DOES VOCABULARY KNOWLEDGE INFLUENCE SPEECH RECOGNITION IN ADVERSE LISTENING CONDITIONS?

A thesis submitted in partial fulfilment of the

requirements for the Degree of

Master of Audiology

in the Department of Communication Disorders

at the University of Canterbury

by Joseph Dalrymple-Alford

2014

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my primary supervisor Megan McAuliffe and secondary supervisor Don Sinex for all their patient assistance and encouragement. I would also like to express my gratitude to my family, friends and classmates for their support.

Abstract

Purpose: To investigate the effects of vocabulary, working memory, age, semantic context, and signal-to-noise-ratio (SNR) on speech recognition in adverse conditions (multitalker babble) in normal-hearing listeners aged 18-35. First, a general hypothesis was tested that listeners with larger receptive vocabularies would be more accurate at recognising speech in noise than listeners with more limited receptive vocabularies, even when target stimuli are words with high lexical frequency. A second more specific hypothesis was that the vocabulary would be predictive of speech recognition accuracy when the signal was moderately degraded, but not mildly or severely degraded.

Method: 80 sentences with a high (HP) or low (LP) degree of semantic predictability (40 HP and 40 LP) were recorded from a male speaker of NZ English. These sentences were used as experimental target stimuli, and presented in multitalker babble at four SNRs: -8, -4, 0 and 4 dB SNR. Thirty-five participants (11 males and 24 females, aged 18 to 35), with puretone hearing thresholds of 15 dB HL or better, completed the Peabody Picture Vocabulary Test (PPVT) and the Wechsler Adult Intelligence Scale (WAIS) vocabulary subtest, the WAIS working memory subtests, and the experimental listening task in which they were required to repeat back the target sentences.

Results: There was considerable variability between listeners in speech recognition performance, in terms of percent words accurately recognised overall (*M* = 45.8%; SD = 7.4) and for both HP (*M* = 54.4%; SD = 9.8) and LP (*M* = 35%; SD = 8.9) conditions. Hierarchical multiple regression analyses revealed that receptive (PPVT) and productive (WAIS) vocabulary knowledge, but not working memory, contributed significant variance to listeners’ speech recognition scores overall and in both the HP and LP conditions. Further regression analyses at individual SNR levels showed that receptive vocabulary contributed significant variance to listening recognition scores in all predictability and SNR conditions except the most favourable (HP stimuli at 4 dB SNR) and least favourable (LP stimuli at -8 dB SNR) listening conditions. Working memory and age were not significantly related to overall listening score, HP listening score, or LP listening score, but age did contribute significant variance in the - 4dB SNR LP condition.

Conclusion: The results provide further evidence that greater vocabulary knowledge is associated with improved speech recognition in adverse conditions. This effect was salient in mid-range adverse listening conditions, but was not apparent in highly favourable and extremely poor listening conditions. The results were interpreted to suggest that in moderately adverse listening conditions listeners with larger lexicons may be better able to exploit redundancies and/or intelligible ‘glimpses’ in the speech signal.

1. Introduction

There is considerable variability across listeners in speech recognition proficiency in adverse listening conditions, even among young normal-hearing listeners (Mattys, Davis, Bradlow & Scott, 2012; Baskent, 2010). Recently, interest has grown in uncovering the factors that may explain this between-listener variability (Mattys et al, 2012; Gilbert, Tamati & Pisoni, 2013). Cognitive ability is one factor commonly shown to be related to speech recognition in adverse conditions, particularly measures of working memory and speed of processing (Akeroyd, 2008). However, most studies investigating the link between individual differences in cognition and individual differences in speech recognition have employed participants who varied in terms of age and hearing acuity. Considered overall, those studies have found that while cognition does influence speech recognition, its role is secondary to that of hearing acuity (Akeroyd, 2008).

More recent studies focussing on underlying cognitive characteristics that may influence listening proficiency in adverse conditions have controlled for the effect of hearing acuity and reported a link between vocabulary knowledge and speech recognition (McAuliffe, Gibson, Kerr, Anderson & LaShell, 2013; Benard, Mensink & Baskent, 2014; Tamati, Gilbert & Pisoni, 2013; Janse & Adank, 2012). However, other studies have not found this association (Benichov, Tun & Wingfield, 2012; Jerger, Jerger & Pirozzolo, 1991).

While the literature regarding the influence of vocabulary knowledge on listening in adverse conditions is somewhat equivocal, it is also incomplete. The existing literature has not addressed whether the degree of adversity in listening conditions influences the effect of vocabulary on speech recognition, nor has it specifically addressed the role of semantic and contextual predictability. Further research in these areas could enhance our understanding of how and why vocabulary knowledge is related to speech recognition accuracy in adverse conditions.

Therefore, the current study had two main aims. Firstly, the current study aimed to add to prior work from this laboratory (McAuliffe et al., 2013) and others (Benard et al, 2014; Tamati et al, 2013; Janse & Adank, 2012) to confirm the effect of vocabulary knowledge on speech recognition. In achieving this aim, the study addressed some potential reasons for inconsistency in the existing literature. Secondly, the current study aimed to extend the existing literature by examining the role of vocabulary knowledge in greater detail than has been reported thus far. Therefore, the study considered the hypothesis that listeners with larger vocabularies have superior top-down processing of speech by examining the influence of vocabulary on speech recognition in varying degrees of adverse conditions, and also in both high and low levels of semantic and contextual predictability.

The following literature review will outline what adverse listening conditions are, before reviewing research on individual variability in speech recognition in adverse listening conditions that has focussed on the influence of cognition and hearing acuity. Subsequently, the review will focus specifically on studies that investigated the role of vocabulary in listening in adverse conditions. Finally, the aims and hypotheses of the current study will be outlined.

1.1 Listening in adverse conditions

Adverse listening conditions may be defined as “any factor leading to a decrease in speech intelligibility on a given task relative to the level of intelligibility when the same task is performed in optimal conditions” (Mattys, Davis, Bradlow & Scott, 2012, p. 953). Considering the listening tasks we engage in daily it is apparent that, for many people, the majority of listening takes place in suboptimal and, therefore, adverse conditions (Assmann & Summerfield, 2004). Much of our listening is performed in environments with competing noise, while our attention is divided between more than one task, or when the signal is artificially altered - such as with the limited bandwidth signal from a telephone (Mattys, White & Melhorn, 2005; Assmann & Summerfield, 2004). Hence, optimal conditions are the exception rather than the rule.

Despite the stresses placed upon speech communication by adverse conditions, normal hearing listeners are usually able to communicate successfully (Assmann & Summerfield, 2004). Listeners are able to understand speech in the presence of significant background noise and in reverberant environments, understand talkers with varying accents, and communicate when speech is distorted by devices such as hearing aids or cochlear implants. Research has confirmed that recognition of speech can be obtained despite severe artificial distortion of the speech signal. For example, when the formants of speech are re-synthesised as sinusoids, the signal remains comprehensible (Remez, Rubin, Pisoni, & Carrell, 1981). Assmann and Summerfield contend that the general success of communication in adverse conditions is due to the robustness of speech’s multiple levels of redundancy, including acoustic, phonetic and linguistic redundancy.

Deployment of compensatory top-down processing resources also assists in listening in adverse conditions (Wild, Davis & Johnsrude, 2012).

Because adverse listening conditions are the norm rather than an exception, Mattys et al. (2012) argue, a general theory of speech perception must be inclusive of listening in adverse conditions rather than treating it as a special case. Mattys et al (2012, p. 963) claim that “speech recognition under adverse conditions is, by and large, synonymous with speech recognition”.

There are a variety of adverse listening conditions, which Mattys et al (2012) classify adverse conditions according to their causes and effects. The effects of adverse conditions on the listener can include reduced attentional or memory capacity, failure to recognise the speech, and perceptual interference. Classified by cause, one example of adverse conditions are those due to *source degradation*, when the speech signal is degraded relative to an exemplar produced by healthy native speakers including accented speech and disordered speech such as dysarthria (Mattys et al., 2012). Adverse conditions resulting from degradation occurring in the transfer of the signal from the speaker to the listener fall into the *environmental/transmission degradation* category. The most commonly encountered example is interference from noise in the environment, while reverberation and distortion produced by a device such as a telephone or a hearing aid are other examples (DeConde-Johnson, 2009; Valente & Valente, 2009). Adverse conditions resulting from *limitations of the perceptual or cognitive abilities of the listener* may be due to a peripheral hearing loss or an incomplete language model such as that of a non-native speaker (Mattys et al, 2012).

Just as adverse listening conditions vary, the ability of individual listeners to cope with a degraded speech signal also shows significant variability. Adverse conditions are ideal for studying individual variability in speech perception, because there is much less variability between individuals in terms of listening proficiency when conditions are ideal (Tamati et al., 2013). Hence, employing adverse conditions to induce greater performance variability may aid our understanding of how and why discrepancies occur (Gilbert et al, 2013). The reasons for variability in speech recognition performance are not clear, but interest in exploring individual differences which may explain the between-listener disparity has focussed on two factors: hearing acuity and cognition.

1.2 Individual differences in speech recognition proficiency in adverse conditions

Although listeners are able to overcome large variations and distortions in the speech signal to recognise speech (Luce & McLennan, 2005), even those with normal hearing acuity vary a great deal in their proficiency at recognising speech in adverse listening conditions (Mattys et al., 2012; Baskent, 2010). That is, with an identical target signal and an identical degree of adverse listening conditions, some listeners are more accurate at recognising speech than others (Gilbert et al., 2013). For example, listeners display substantial variability in the degree to which their listening accuracy is affected by competing signals (e.g. Wightman, Kistler & O'Bryan, 2010; Gilbert et al., 2013) or interruptions in the signal (Baskent, 2010), in their proficiency in adapting to foreign accents (Janse & Adank, 2012) and their comprehension of regional dialects (Mason, 1946; Gilbert et al., 2013).

Hearing acuity and cognition are two factors that have been proposed to account for variability in speech recognition accuracy. Clearly, speech recognition in adverse conditions will be influenced by hearing acuity (McArdle & Hnath-Chisolm, 2009). A large body of research has confirmed this, especially in the case of sensorineural hearing loss, which affects not only audibility but also temporal and frequency resolution (Dillon, 2012). When listening to speech in competing multitalker babble, a listener with a hearing loss is likely to require an SNR 10-15 dB higher than a normal-hearing listener to achieve the same level of speech recognition accuracy (e.g. Olsen & Carhart, 1967; Carhart & Tilman, 1970). The speech perception of listeners with hearing loss is also more significantly affected than normal-hearing listeners by reverberation (Olsen, 1981) and distance (Smaldino, Crandell, Kreisman, John & Kreisman, 2009).

Given that speech recognition variability in adverse conditions exists even among normal-hearing listeners, however, hearing loss alone cannot account for the trend. An important factor that has been cited as an influence is cognition. In a challenging listening environment, the listener is tasked with making sense of a signal that is degraded or ambiguous, a situation likely to result in much greater allocation of top-down processing resources (Baskent, 2010; Warren, 1984). Therefore, it seems likely that an individual whose fluid cognitive abilities are intact and capable would be at an advantage when listening conditions are difficult (Schneider, Li & Daneman, 2007). Such an individual would, for example, be better able to employ their working memory to keep track of relevant information, process information at a more rapid rate so that it may be integrated with previous information, or better infer the likely identity of an indistinct word. Hence, the finding that listeners vary in their speech recognition proficiency in adverse conditions has led to numerous studies investigating whether cognitive factors could explain the variation, some of which will be outlined in the following section.

In general, studies have reported that cognition does account for individual variability in speech recognition in adverse listening conditions (Beer, Kronenberger & Pisoni, 2010; Gordon-Salant & Fitzgibbon, 1997; Mattys et al, 2012), but only after the much stronger influence of hearing acuity has been accounted for (Akeroyd, 2008; Humes, 2002).

A review article by Akeroyd (2008) described 20 experimental studies which had investigated the relationship between individual differences in cognition and variability in speech recognition accuracy in noise. Most of the studies surveyed included participants with a range of age and hearing acuity. In 19 of the 20 studies reviewed, some aspect of cognition was significantly related to speech recognition performance in noise (Akeroyd, 2008). However, the effect of cognition was secondary to the effect of hearing acuity in all of the studies that included participants with hearing loss. Furthermore, while the results of the twenty studies reviewed demonstrated a link between cognitive function and speech perception, the studies were inconsistent in terms of which aspects of cognition tended to obtain a significant effect, with measures of working memory most often successful (Akeroyd, 2008).

Variation in degree of cognitive decline in older listeners has also been shown to explain some of the variation in speech recognition performance in this population over and above the effect of hearing acuity (Schneider, Daneman & Pichora-Fuller, 2002). For example, Humes (2002) reported speech recognition and WAIS-R results for 171 older listeners who used hearing aids. Humes found that the strongest predictor of listening accuracy was hearing acuity followed by ‘nonverbal IQ and ageing’ and then verbal IQ.

To sum up, hearing acuity has been shown exert a significant influence on individual variability in speech recognition proficiency in adverse conditions. Cognition is also related to the individual variability in speech recognition, but is a much weaker predictor of proficiency in the presence of a hearing loss. However, most studies which have investigated individual differences in speech perception in adverse conditions have included participants with a range of hearing acuity. Hence, it is difficult to determine which cognitive factors play a role in speech perception. More studies are thus required which control for hearing by including only participants with normal hearing acuity. Recently, a number of such studies have been published that have shed light on a factor that had not been previously considered: vocabulary knowledge.

1.3 The influence of vocabulary knowledge on listening in adverse conditions

Earlier research on the relationship between vocabulary knowledge and speech recognition focussed on listening difficulty, second language learners and children (Howes, 1957; Munson, 2001; Bradlow & Pisoni, 1999). The idea that vocabulary knowledge could influence speech recognition in normal-hearing adults, even when the target speech comprises words with high lexical frequency, is a very recent one. In the last few years, a handful of research groups have specifically considered the hypothesis that individual differences in vocabulary or linguistic skill could explain some of the variability in speech recognition proficiency in adverse listening conditions. The results of these studies have mostly favoured this hypothesis, finding an influence of vocabulary knowledge on speech perception with disordered speech (McAuliffe et al, 2013), adaptation to accented speech (Janse & Adank, 2012), interrupted speech (Benard, Mensink & Baskent, 2014) and speech in noise (Tamati et al, 2013), whereas Benichov, Tun and Wingfield (2012) and Jerger Jerger and Pirozzolo (1991) have reported conflicting findings. The following section will first briefly review earlier research which investigated the relationship between vocabulary knowledge and speech recognition from a more general perspective. Subsequently, the studies which have explicitly investigated the relationship between vocabulary knowledge and speech recognition in adverse listening conditions will be reviewed in greater detail.

Lexical factors have been shown to affect the difficulty of listening tasks. For example, high lexical neighbourhood density has been demonstrated to reduce the speed and accuracy of spoken word recognition (Altieri, Gruenenfelder & Pisoni, 2010) and lexical frequency and familiarity affect speech recognition in noise (Howes, 1957; Pollack, Rubenstein & Decker, 1959). Furthermore, vocabulary knowledge has been shown to relate to second language speech recognition (Bradlow & Pisoni, 1999; Stæhr, 2009). Stæhr (2009) reported that listening comprehension performance was strongly correlated with receptive vocabulary in the sample of 115 highly proficient English as a second language learners. Vocabulary has also been shown to influence child language development (Munson, 2001; Edwards, Beckman & Munson, 2004). Munson (2001) found that expressive and receptive vocabulary measures could explain a significant portion of the variation in speech recognition scores in children aged 3 – 7. Additionally, there is speculation that older adults may be superior to younger adults at using semantic context to boost speech recognition in adverse listening conditions (Pichora-Fuller, Schneider &

Daneman 1995; Pichora-Fuller & Souza, 2003), because older adults tend to have a larger vocabulary than their younger counterparts (Sheldon et al, 2008).

To the author's knowledge, only the six studies earlier mentioned have specifically examined the link between vocabulary knowledge and speech recognition proficiency in adverse conditions. Four found a positive relationship (McAuliffe et al, 2013; Benard et al, 2014; Tamati et al, 2013; Janse & Adank, 2012), but two reported no relationship existed (Benichov et al, 2012; Jerger et al, 1991). The reason for this inconsistency is not yet known, but all six studies differ in terms of the adverse conditions induced, the vocabulary assessments employed, the stimuli used, and the degree to which hearing acuity was controlled for.

The four studies that found a relationship suggest that the influence of vocabulary on speech recognition is robust across different types of adverse conditions, as each has used different stimuli and a different method of degrading the speech signal. The current study was designed to follow on from a study conducted in the same research laboratory by McAuliffe et al (2013), who induced adverse conditions by employing dysarthric speech as the target stimuli. Hypokinetic dysarthric speech is characterised by monopitch, monoloudness, phoneme imprecision and a fast rate of speech (Darley, Aronson & Brown, 1969). The adverse conditions were further intensified by using phrasal stimuli which offered no lexical or contextual predictability to aid recognition. Benard et al (2014), by contrast, employed high context phrases which were interrupted with either silence or bursts of noise and presented at two speech rates, while the adverse conditions in the study conducted by Janse and Adank (2012) involved listening to an unfamiliar accent. The adverse conditions in Tamati et al (2013) involved listening to The Perceptually Robust English Sentence Test Open-set (PRESTO; Gilbert et al, 2013), with phrases spoken by multiple talkers in a variety of regional accents, with additional variability added by various levels of syntactical complexity and SNR (Gilbert et al., 2013; Tamati et al, 2013). Overall, these studies suggest the relationship between vocabulary knowledge and speech recognition in adverse conditions does not depend on the type of adverse conditions or the presence or absence of semantic or contextual predictability in the target phrases.

Of the four studies which found a positive relationship between speech recognition and vocabulary, two included only participants with normal hearing thresholds of 20 dB HL or better (Benard et al, 2014; Tamati et al, 2013), one included participants with normal hearing acuity or a slight hearing loss (=/< 25 dB HL; McAuliffe et al, 2013) and one included participants with hearing acuity ranging from normal to moderate hearing loss (Janse & Adank, 2012). McAuliffe et al (2013) reported that for younger participants, vocabulary predicted listening recognition accuracy, whereas in the older group no main effect of vocabulary or hearing existed, but an interaction between vocabulary and hearing was present. This interaction demonstrated that higher vocabulary scores predicted higher listening accuracy when hearing acuity was normal for older participants, but this effect was mitigated by elevated hearing thresholds in the older participants, despite the fact that the worst-hearing older participants had only slightly elevated thresholds of 25 dB HL. Janse and Adank found that hearing acuity predicted overall listening performance, but vocabulary did not. Rather, vocabulary was related to adaptation to the adverse conditions which involved listening to an unfamiliar accent. Together the results of McAuliffe et al (2013) and Janse and Adank (2012) suggest that elevated hearing thresholds could diminish or eliminate the effect of vocabulary on speech recognition proficiency.

With regard to age, it appears that vocabulary can influence speech perception in both younger and older listeners. Benard et al (2014) and Tamati et al (2013) included only younger adults, whereas Janse and Adank (2012) only older adults and McAuliffe et al (2013) included both younger and older adults, all finding that vocabulary was positively related to speech perception in adverse conditions. As mentioned, working memory is often linked to speech recognition proficiency in adverse conditions (Akeroyd, 2008). However, the results of the four studies with regard to the influence of working memory are inconsistent, with two studies reporting an association between working memory and speech recognition (Janse & Adank, 2012; Tamati et al, 2013) and two reporting no association (McAuliffe et al, 2013; Benard et al, 2014).

Overall, the four studies demonstrated that a positive relationship between vocabulary and speech recognition exists in normal-hearing listeners in various types of adverse conditions, with both high and low context stimuli, for both older and younger adults. The studies also suggest that hearing loss negates the effect of vocabulary knowledge, and are inconsistent in their conclusions regarding working memory.

Not all studies which have looked at vocabulary's role in variability in listening in adverse conditions have found a relationship between a larger lexicon and superior listening proficiency, however. Benichov, Tun and Wingfield (2012) and Jerger, Jerger and Pirozzolo (1991) both did not find an association. The two studies share similarities: both employed participants with a range of hearing acuity, both used speech masked by noise as the adverse condtions, and both used the WAIS vocabulary subtest to assess vocabulary knowledge. Benichov, et al (2012) did not find a relationship between vocabulary and speech recognition in the adverse conditions induced with background multitalker babble and various levels of contextual and semantic predictability. The results showed that hearing acuity was a strong predictor of speech recognition performance for low context stimuli, but that as the degree of context increased, the influence of hearing acuity was reduced and a composite cognitive ability (comprised of working memory, episodic memory and speed of processing) became a more important predictor of task performance, while vocabulary was not significantly associated with speech recognition (Benichov et al, 2012)

The second study that did not find an influence of vocabulary knowledge on speech recognition in adverse conditions was conducted by Jerger et al (1991), who carried out a correlational analysis of speech recognition, hearing loss, age and a range of cognitive abilities including vocabulary. The 200 participants (*M* age = 70; average puretone hearing threshold (PTA) = 32 dB HL) were given a battery of neuropsychological tests including all subtests of the WAIS-R. Hearing acuity accounted for the majority of the variance in the listening recognition performance, with the combined cognitive variables adding only 3 to 6% of variance. No single cognitive measure, including vocabulary, significantly incremented prediction (Jerger et al., 1991).

Why is there mixed evidence regarding the influence of vocabulary knowledge on speech recognition? There are a number of possibilities. Firstly, the studies mentioned in the above review used different measures of vocabulary, including the Peabody Picture Vocabulary Test (PPVT; used by McAuliffe et al, 2013 and Benard et al, 2014), a receptive word familiarity task (Tamati et al, 2013), a computer-based multiple-choice test of receptive vocabulary (used by Janse & Adank, 2012), and the vocabulary subtest from the Wechsler Adult Intelligence Scale (WAIS; used by both

Benichov et al, 2012 and Jerger et al, 1991). It is interesting to note that both studies which did not find a significant effect of vocabulary size on speech understanding in older adults used the same measure: the vocabulary subtest from the WAIS, which is a productive test of vocabulary, rather than a receptive test such as those employed by McAuliffe et al (2013), Tamati et al (2013), Benard et al (2014) and Janse & Adank (2012). Therefore, it is possible that the measure itself could explain the inconsistency in the findings between studies.

A second possibility for the inconsistency in the findings is variation in the degree of complexity between the tasks used in the studies. Benichov et al and Jerger et al both employed SPIN sentences in which the task was to repeat the final word of the sentence presented, whereas the task in the study of McAuliffe et al (2013) and Tamati et al (2013) was to repeat the entire sentence. Repeating the complete sentence is a more challenging task which could lead to greater involvement of vocabulary knowledge in parsing and recognising the words.

A third potential reason for the conflicting findings is hearing acuity. The studies conducted by Benard et al, McAuliffe et al and Tamati et al included normal-hearing or near normal-hearing participants, whereas Benichov et al and Jerger et al included participants with a range of hearing acuity. In both Benichov et al and Jerger et al, hearing acuity was found to be the most important predictor of speech recognition performance. In McAuliffe et al, hearing acuity was found to moderate the effect of vocabulary on listening proficiency in the older participants, despite the inclusion of participants with hearing thresholds of no greater than 25 dB HL. Given this result, it seems possible that any effect of vocabulary was overridden by hearing acuity in the studies of Benichov et al and Jerger et al.

The preceding section has reviewed the existing research on the relationship between vocabulary knowledge and speech recognition in adverse conditions and addressed potential reasons for inconsistency in the literature. While acknowledging that the literature remains equivocal, the following section will consider how vocabulary knowledge might relate to top-down linguistic processing which could improve speech recognition in adverse listening conditions.

1.4 Reasons for the relationship between vocabulary knowledge and speech recognition in adverse conditions

The reason for a relationship between vocabulary and speech recognition in adverse conditions is not immediately obvious, but if listeners with larger vocabularies do not simply know more of the target words, they may instead have superior top-down processing of degraded speech (McAuliffe et al, 2013).

McAuliffe et al (2013) found that the lexical frequency of the items was not related to the listening accuracy, confirming that the effect of vocabulary found was not due to greater familiarity of participants with the particular stimulus words employed. Instead, McAuliffe et al raised the possibility that individuals with larger lexicons were better able to make use of the intelligible ‘glimpses’ within the hypokinetic dysarthric speech, due to their assumed greater familiarity and experience with language in general. This greater familiarity and experience with language, McAuliffe et al proposed, meant listeners with higher vocabulary scores “were better able to make use of redundancies within the acoustic signal and, ultimately, leverage this prior experience to draw accurate lexical hypotheses” (McAuliffe et al., 2013, p. 1366).

The notion that listeners can use glimpses of the target signal to segregate speech was first proposed by Miller and Licklider (1950), who assessed the intelligibility of speech that was interrupted at a range of modulation frequencies. It was found that high levels of speech understanding remain at an interruption rate of 10 Hz, even though half of the signal is absent at this modulation frequency. Miller and Licklider concluded that listeners were able to piece together glimpses of the target speech which are accessible during the uninterrupted fragments of speech. Further evidence of glimpsing was reported by Festen & Plomp (1990), who confirmed that listening recognition performance is lower in steady-state noise than in modulated noise or a single competing talker, suggesting listeners exploit silent gaps in the competing signal to recognise the target words (Festen & Plomp, 1990).

The ‘listening-in-the-gaps’ theory of speech segregation is not feasible when there are four or more competing voices present, however, because the masking waveform is nearly continuous (Miller, 1947). Cooke (2003; 2005) proposed a different definition of what constitutes a ‘glimpse’: “a time-frequency region which contains a reasonably undistorted ‘view’ of local signal properties” (Cooke, 2005, p. 306). In this view, the strategy listeners confronted with speech in noise may employ is to first search within the noise for regions (glimpses) which could signify fragments of speech (Cooke, 2006). Recognition of the speech signal is then based on the incomplete information provided by the glimpses, with the burden of recovering the target speech placed on top-down processing resources.

Support for the notion that top-down influences are important in recognising degraded speech has also come from functional magnetic resonance imaging (fMRI) studies. For example, Wild, Yusuf, Wilson, Peele, Davis and Johnsrude (2012) demonstrated that the comprehension of speech in adverse listening conditions and the engagement of brain areas that support speech processing depend on attention. Wild et al (2012) showed that attention enhances the processing of degraded speech by engaging higher-order top-down mechanisms and found a pattern of response which they suggested represents the neural correlate of effortful listening. These results are consistent with other fMRI studies which have found that speech recognition in adverse listening conditions is facilitated by top-down influences on auditory processing (Davis and Johnsrude, 2007; Sohoglu, Peele, Carlyon & Davis, 2012).

Overall, the literature indicates that top-down processing is important in listening in adverse conditions, and listeners are likely to exploit redundancies and intelligible fragments in the signal to assist them in recognising degraded speech.

If the speculation that listeners with larger vocabularies tend to have superior top-down linguistic processing is correct, we might expect the following pattern of results with regard to the degree of degradation in the signal: (1) when listening conditions are good, vocabulary does *not* influence speech recognition because the signal is rich, top-down influence is negligible, and minimal variability between listeners exists, (2) when listening conditions are moderately adverse, vocabulary *does* influence speech recognition because listeners with larger vocabularies are better able to exploit intelligible glimpses and leverage redundancies in the speech signal, (3) when listening conditions are very adverse, vocabulary does *not* influence speech recognition because the speech signal is so degraded that no intelligible glimpses or redundancies remain, leaving listeners to resort to sublexical cues such as syllabic stress in their attempts to recognise the target words (Mattys et al., 2005; 2012).

One way to achieve various levels of adverse listening conditions is to use noise to systematically degrade the speech signal by varying the SNR; a high SNR would constitute less of an adverse condition than a low SNR. Therefore, in addition to an overall relationship between vocabulary size and task listening performance, any interaction between vocabulary ability and listening performance at different degrees of adverse listening conditions could be examined.

A second way to vary the level of adversity in listening conditions is to directly compare different levels of predictability in the target stimuli. High context stimuli are easier to recognise in adverse conditions than low context stimuli (Dubno, Ahlstorm & Horwitz, 1999). Previous studies which have reported an influence of vocabulary knowledge on listening in adverse conditions have used *either* phrases with a low level (McAuliffe et al, 2013) or a high level (Benard et al, 2014) of semantic and contextual predictability. This suggests the effect of vocabulary knowledge remains regardless of the level of predictability. However, because high context stimuli have semantic and contextual cues in addition to phonemic and other sublexical cues, we might expect that listeners with superior vocabulary (and, we assume, superior top-down linguistic processing) would have even more of an advantage with high context stimuli compared to low context stimuli. Therefore, inclusion of speech stimuli with both high and low levels of predictability matched for length and lexical frequency would allow us to examine a differential effect of vocabulary knowledge according to the degree of stimulus predictability.

1.5 Aims and hypotheses of the current study

The current study has two primary aims. Firstly, the study aims to confirm the relationship between vocabulary knowledge and speech recognition by addressing the inconsistencies present in conflicting studies described in a preceding section.

To achieve this aim, the design of the current study differed from previous research in two ways. Firstly, the current study employed *both* a productive vocabulary test (the WAIS IV subtest) and a receptive vocabulary test (the PPVT – 4^{th} Edition), because previous research has tended to use either a receptive measure or a productive measure to assess vocabulary, and these measures have returned differing results. Secondly, because previous studies have shown that hearing acuity influences the effect on speech recognition of cognition generally (Akeroyd, 2008) and vocabulary specifically (McAuliffe et al, 2013), the current study controlled for hearing acuity by setting a conservative inclusion criteria for participants of hearing thresholds of 15 dB HL or better in both ears.

With regard to the first aim, the following hypotheses were formulated:

(1) It was hypothesised that receptive vocabulary knowledge as assessed by the PPVT-IV would be related to speech recognition performance in adverse conditions overall and in both HP and LP conditions. Specifically, it was expected that listeners with larger receptive vocabularies would recognise a larger proportion of the speech in both HP and LP conditions.

(2) Based on previous research, it was hypothesised that productive vocabulary knowledge as assessed by the WAIS vocabulary subtest will not be related to listening performance in adverse conditions.

(3) Because the age range of participants in the current study was relatively narrow and hearing acuity has been controlled for, it was not expected that age would be related to listening recognition performance.

(4) The current study used the short phrases the same as (LP) and similar to (HP) those employed by McAuliffe et al (2013), who found no relation between working memory and speech recognition performance in adverse conditions. Therefore, no relationship between working memory and listening recognition performance is anticipated in the current study.

The second aim of the study was to examine whether listeners with larger vocabularies tend to have superior top-down processing of degraded speech by determining whether the influence of vocabulary on speech recognition varies according to the degree of adversity in listening conditions and the degree of semantic and contextual predictability in the target. Therefore, the study used systematically degraded speech samples with multitalker babble at various SNRs to induce adverse listening conditions of varying degrees and included target stimuli with two levels of semantic and contextual predictability in the stimuli. The same LP sentence stimuli as used by McAuliffe et al were included in the current study, as well as a set of sentences with high semantic and contextual predictability (HP).

With regard to the second aim of the study, the following hypotheses were formulated:

(1) It was anticipated that participants would successfully recognise a greater proportion of the high predictability stimuli than the low predictability stimuli.

(2) It was expected that speech recognition performance will decline as the signal-to-noise ratio deteriorates.

(3) It was hypothesised that participants with larger receptive vocabularies would tend to recognise more of the speech stimuli than participants with smaller receptive vocabularies when listening conditions were *moderately adverse*.

(4) When listening conditions were *favourable* (i.e. high context and high SNR), or *extremely poor* (i.e. low context and low SNR), it was expected that receptive vocabulary knowledge would not be related to listening recognition performance.

(5) It was hypothesised that receptive vocabulary will be more strongly related to speech recognition accuracy in the HP condition than the LP condition.

2. Method

2.1 Participants

Thirty-five participants were recruited for the study. The 11 males and 24 females (*M* age = 24 years) were native speakers of New Zealand English aged between eighteen and thirty-five years. All participants exhibited normal hearing acuity bilaterally as measured via behavioural pure-tone audiometry screening in a soundproof booth. Pure-tone stimuli were presented via supra-aural headphones at 15 dB HL at the audiometric frequencies of 500 Hz, 1 kHz, 2 kHz, 4 kHz and 8 kHz. The usual threshold of ‘normal’ hearing is 20 dB HL or better (Schlauch & Nelson, 2009), so a screening threshold of 15 dB HL was chosen as a conservative measure to ensure that hearing acuity was fully controlled for and would not play a role in listening performance variability between listeners. Initially 38 participants were recruited, however, three participants exhibited behavioural thresholds of greater than 15 dB HL at one or more audiometric frequencies in either ear and were excluded, leaving 35 to complete the study.

The majority of the participants were current students of the University of Canterbury, and the remainder were associates of the principal researcher. Participants were compensated for their involvement in the study. Approval for the study was obtained from the University of Canterbury Human Ethics Committee. All participants were fully informed of the procedures and aims of the study and signed a consent form prior to participating.

2.2 Listening task stimuli

In order to investigate the influence of contextual and semantic predictability on the listening task, two sets of experimental phrases were employed: a set of phrases with a high level of inter-word predictability (HP) and a set with low inter-word predictability (LP). Each set was comprised of 40 six syllable sentences, with the high and low inter- word predictability sets having mean lengths of 5.4 (SD = 0.59) and 4 (SD = 0.68) words, respectively. Additionally, four different HP phrases and four LP phrases were selected to be used as practice stimuli.

The HP set of 40 phrases ranged from four to six words per utterance, and were selected from sentences adapted by McAuliffe, Wilding, Rickard & O'Beirne

(2012) from the Speech in Noise (SPIN) Test stimuli (Kalikow, Stevens & Elliot, 1997). These phrases are designed to contain a high degree of semantic and contextual sentential predictability, such that it was possible to infer the identity of words included in the phrase even if they are not heard or only partially heard. The LP phrases were selected from a set developed by Liss, Spitzer, Caviness, Adler & Edwards (1998), modelled on sentences developed by Cutler and Butterfield (1992). These phrases were designed to reduce the contribution of semantic and contextual information to intelligibility (Liss et al., 1998). The LP phrases ranged from three to five words per phrase, and alternated strong (S) and weak (W) syllables, meaning half of the phrases contained an SWSWSW phrasal stress pattern, and half contained a WSWSWS phrasal stress pattern (Liss et al., 1998). See Table 1 below for examples of HP and LP sentences.

HP stimuli	*LP stimuli*
The man wrote a letter. Soup is served in a bowl. Pour water down the drain. The landlord raised the rent. Pick a bunch of flowers.	Mark a single ladder. May the same pursued it. Attend the trend success. Technique but sent result. Account for who could knock.

Table 1: Examples of high predictability (HP) and low predictability (LP) sentence stimuli.

The Range computer program (Nation & Heatley, 1994; Heatley, Nation & Coxhead, 2002) was used to confirm that the words used in the stimuli sentences were common words which should be familiar to all of the participants. The range program compares the stimulus words against lists derived from the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) (a combined corpora of 550 million words) ranked in order of word frequency. Both HP and LP stimuli lists contained very common words that should be recognised by a native English speaker, with 94.9% of LP words and 96.5% of HP words falling in the 4000 most common words of English according to the BNC and COCA (see Table 2). By way of comparison, one group of researchers estimates that an English speaking university graduate has a vocabulary of approximately 20,000 words (Goulden, Nation & Read, 1990). Therefore, it is likely that all of the stimulus words fell within the lexicons of the participants.

Ranking according to BNC and COCA	**HP**	**LP**
1000 most common words in English	60.28%	60.58%
2000 most common words in English	78.72%	81.75%
3000 most common words in English	90.78%	87.59%
4000 most common words in English	96.45%	94.89%
6000 most common words in English	100%	97.08%

Table 2: Cumulative proportion of low predictability (LP) and high predictability (HP) stimuli words in terms of word frequency based on the BNC and COCA corpora.

2.2.1 Recording and preparation of speech stimuli

The speech stimuli used in the study were recorded from a 32-year-old male native speaker of New Zealand English who was compensated for his participation. The speaker had no history of neurological disorder, speech or language disorder, or hearing loss, and did not have a temporary condition such as an upper respiratory tract infection which could have affected his speech production. The speech samples were collected during a single session of approximately one hour.

The speaker was requested to read lists of 44 LP phrases and 44 HP phrases in his “normal conversational voice”. The speaker was allocated five minutes to read and familiarise himself with typed lists of the phrases. An Audix HT2 Headset Condenser Microphone was then positioned with the microphone approximately five centimetres from the mouth. The phrases were recorded at a sampling rate of 44.1 kHz with 16 bits of quantization.

During the recording, the principal investigator presented cards to the speaker with a single phrase printed in 20 point typeface which the speaker read twice in succession. The speaker was provided with drinking water and allowed as many breaks as necessary throughout the recording process. Once the recording was completed, the principal researcher listened to the recording and selected one of the two readings of each phrase to be included in the experimental stimuli based on fluency and/or lack of mispronunciations or pauses.

Subsequently, the sound files containing the recording were edited with Adobe Audition (Version 3) to eliminate microphone noise and produce individual Wave (.wav) files each containing one phrase with a brief period of silence preceding and following the speech sample. These files were later mixed with the multi-talker babble to achieve the various levels of signal degradation employed.

2.3 Procedure

All participants completed all of the experimental tasks in a single session of approximately one hour duration. The principal researcher conducted all experimentation. All participants completed four tasks in the following order: (1) the Weschler Adult Intelligence Scale (WAIS) working memory digit span task, (2) the Peabody Picture Vocabulary Test (PPVT) receptive vocabulary task, (3) the WAIS productive vocabulary task, and (4) an experimental listening task. The listening task was conducted in a sound-treated booth and the other tasks were completed in an adjoining office in the Speech and Hearing Clinic of the University of Canterbury.

2.3.1 Working memory task

The WAIS IV (2007) digit span subtest provided a well-validated measure of working memory. This test includes three subtasks: a forward digit span, a backward digit span and a sequencing digit span task. In the forward digit span task, the tester read two examples and then read increasingly longer sequences of numbers at a rate of one per second without repetition, and the participant was asked to repeat the numbers back in the same order. The task ended either when the participant made errors on two consecutive trials or all of the trials were completed. The backward digit span and sequencing digit span tasks followed the same protocol except that in the backward digit span task the participant was required to repeat back the numbers in the reverse order from that given by the tester, and in the sequencing digit span task the participant was required to repeat back the numbers in order from lowest to highest. The task was scored individually for each subtask, and each score was converted to a standard score using the tables provided in the WAIS manual. An average of the three standard scores was then derived to give a working memory score for each participant.

2.3.2 PPVT receptive vocabulary task

The Peabody Picture Vocabulary Task, fourth edition (PPVT-IV) (Dunn & Dunn, 2007) is a reliable and well-validated tool for measuring the receptive vocabulary of children and adults which was age normalised on a representative sample of 3,540 individuals in the United States. The current study utilised Form A of the PPVT-IV, which contains 228 items grouped in 19 sets of 12 items.

The test employs an administration easel, with each leaf bearing four pictures, one of which represents the test item. For adult examinees, such as the participants in the current study, the pictorial representation they believe best illustrates the item spoken by the tester is selected by indicating the corresponding number (from 1 – 4) printed in the corner of the picture. The tester then turns the leaf to reveal the next four pictures. The test continues until the participant incorrectly identifies 8 or more items from a set of 12. This set becomes the ceiling set, and the participant's raw score is calculated by summing the number of errors and subtracting this amount from the number of the ceiling item, which is the final item in the ceiling set. The test took between ten and fifteen minutes to complete. Raw scores were converted to standard scores according to the age-referenced standard score charts in Appendix A of the PPVT-IV manual.

2.3.3 WAIS productive vocabulary task

The WAIS (IV) (Weschler, 2008) vocabulary subtest was included in the test battery in order to investigate the relative relationship between productive as opposed to receptive vocabulary and listening task performance.

The WAIS-IV vocabulary subtest is a verbal test of productive vocabulary. For adult participants, the test includes 26 items with two ‘teaching items’ that are all presented orally and visually. The tester informed the participant they would be presented with words, which they were required to “tell me in words what it means”. The tester read each item to the participant and pointed to the word in the WAIS stimulus book as it was pronounced. The participant’s response was manually recorded verbatim on the scoring sheet by the tester. The response was scored either 0, 1 or 2 points according to the comprehensive list of sample answers given in the scoring manual. If the examiner’s response was unclear or vague or followed by a (Q) in the scoring manual, the tester queried the response by saying “Tell me more about it” or “What do you mean?” The test was terminated if the participant’s responses returned scores of 0 on three consecutive trials. Raw scores were converted to standard scores using the WAIS IV manual.

2.3.4 Experimental listening task

The listening task was conducted with the participant seated in a sound proof booth and the tester in the adjoining control room. The phrases had been recorded onto an Asus U43JC laptop computer as wave (.wav) sound files at 1058 kbps bit rate. The computer was connected to the external device port of a Gradson-Stadler GSI 61 two-channel audiometer via RCA leads. Prior to each test session the laptop was calibrated by manually adjusting a calibration tone to 0 dB on the VU meter dial of the audiometer. The tester adjusted the audiometer’s talkback dial so that the participant’s speech was clearly picked up by the microphone in the test booth and could be heard effortlessly. The participants listened to the phrases diotically via Telephonics TDH-SDP supra-aural headphones.

Firstly, four practice phrases were presented in the absence of competing babble at 60 dB HL. The participant was informed that this was the voice that they needed to listen for and that they should attempt to ignore the other voices which would be present in the subsequent experimental task. The participant was asked to attempt to repeat back each phrase exactly as they heard it. They were encouraged to guess if they were not sure, and to give a partial answer if they only heard part of the phrase. They were also asked whether the volume was comfortable and clearly audible and given an opportunity to adjust it if necessary. Most participants were satisfied with the listening level, but two asked for it to be lowered slightly, and heard the remainder of the phrases at 55 dB HL. Once the participant was comfortable with the task and had demonstrated they could repeat back the phrases verbatim in the absence of competing babble, the tester indicated that the experimental phase would begin, and that the subsequent task was identical except for the addition of rival babble which they should attempt to ignore.

The phrases were played via custom software produced using the MATLAB program (The Mathworks, Inc., 2012). The software was used to generate four random sequences of 40 phrases for both the HP and LP phrase lists, with each of the participants hearing one of the four randomly generated sequences. Each participant heard eighty phrases (forty HP and forty LP) in addition to the four practice phrases. Half of the participants heard the forty HP phrases followed by the forty LP phrases, and half the participants heard the LP phrases followed by the HP phrases. Each participant heard twenty (ten HP and ten LP) of the eighty phrases at each of four signal-to-noise (SNR) ratios: -8, -4, 0 and +4 dB SNR. The SNR levels employed were selected in order to provide enough data points to interpolate a performance-intensity function for each participant. It was believed that most participants would recognise close to 0% of the speech at -8 dB SNR and close to 100% of the speech at 4 dB SNR, with recognition scores for -4 and 0 dB SNR stimuli falling somewhere between. Therefore, the listeners would be subjected to a range of degrees of adverse listening conditions, from mildly adverse to extremely adverse. The program randomly assigned the order of the four dB single to noise ratios for each participant and for each phrasal set from one of twenty-four possibilities (-8, -4, 0, 4; -8, 0, -4, 4; 4, 0, -8, -4 et cetera).

Three-talker babble was used as the masking stimulus for the experimental listening task. The babble was produced by overlaying utterances from three speakers (two male, one female) from the GRID Corpus (Cooke, Barker, Cunningham & Shao, 2006). The GRID corpus is a large audiovisual corpus of one thousand sentences collected from 34 native speakers of British English in the form: <verb> + <colour> + <preposition> + <letter> + <digit> + <adverb>. For example: ‘place red by G6 now’ and ‘set green at A4 please’. The current study used audio-only samples of utterances collected from speakers three, four and five from the GRID corpus. Raw audio-only 50 kHz files were downloaded from *http://spandh.dcs.shef.ac.uk/gridcorpus* and rescaled in MATLAB to 44.1 kHz to match the experimental stimuli. The masking babble was produced and mixed with the experimental phrases to be played to the participant in real time by the custom MATLAB software used to present the stimuli. Using the program to do this in real time meant it was not necessary to pre-record each phrase at each of the four signal-to-noise ratios applied.

Each trial was initiated by the tester pressing the space bar on the laptop, which presented one experimental phrase in the presence of the background babble to the participant. The participant was requested to verbally repeat the phrase exactly as they had heard it. The tester manually transcribed their response and read it back to them to ensure it had been accurately recorded. Guessing was encouraged if a participant was not sure of their response, and partial responses were encouraged if a participant had only heard part of the phrase. Once both the participant and the tester were satisfied their response had been correctly transcribed, the next trial was presented. Once ten trials at a particular SNR level had been completed, the software automatically switched to the next SNR ratio in the randomly generated sequence, with both tester and participant blinded to whether it would become easier or more difficult for the subsequent set of ten trials. Once all forty trials of one of the predictability sets were completed, the program automatically switched to the remaining set.

2.4 Scoring

The thirty-five sets of phrase transcriptions were scored by the primary researcher. Each listener response was scored by calculating the number of words accurately recognised according to established procedures (Borrie, McAuliffe & Liss, 2012; Liss et al., 1998). A response was scored correct if it matched the target word, if it was a homonym, if it added or subtracted the “ed” tense ending or the plural “s”, or substituted “a” for “the” or vice versa.

2.5 Statistical Analysis

To analyse the listening recognition task results, the listeners' scores in percent words accurately recognised were first calculated for HP and LP phrases at each SNR. These scores were then converted into radians via the rationalized arcsine transform (Studebaker, 1985). Investigations of speech recognition performance commonly employ an arcsine transform to convert test scores in percent into radians for evaluation. This process "adjusts the dimensions of the percentage scale so that: (1) scores in all parts of the scale except the ends are normally distributed around their mean values; (2) mean scores and variances are not correlated with one another; and (3) the likelihood that a score will increase or decrease remains constant over most of the performance range" (Sherbecoe & Studebaker, 2004, p. 442). Performing the rationalized arcsine transform assists in complying with the assumptions of statistical procedures used to analyze the listening recognition scores and the derived scores have the additional advantage of resembling the percentages they represent, making intuitive interpretation easier (Studebaker, 1985).

Subsequently, a two-way repeated measures ANOVA was conducted to examine the main effects of SNR and predictability on RAU speech recognition scores, and any interaction effect between them.

The RAU speech recognition scores were then analysed via regression analysis. A series of correlations were calculated between RAU speech recognition scores and participant variables including PPVT vocabulary score, WAIS vocabulary score, WAIS working memory score and age. The predictor variables were chosen on the basis that all had been shown to be associated with listening in adverse conditions in previous investigations (McAuliffe et al, 2013; Benichov et al, 2012). Vocabulary was chosen as the first variable to be entered into the hierarchical model due to the findings of McAuliffe et al (2013), of which the current study is a partial replication and which used the same LP stimuli, that vocabulary knowledge was related to listening in adverse conditions. Secondly, working memory was included because a number of studies have found a relationship between working memory and listening in adverse conditions. Age was entered third as prior studies had tended to find an effect of age on listening in adverse conditions

In total, eleven hierarchical multiple regression analyses were carried out: for overall scores with all levels of SNR and context, for each level of context (HP and LP) including all levels of SNR, and separately for each level of SNR (LP-8; LP-4; LP0; LP4; HP-8; HP-4; HP0; HP4). Separate hierarchical multiple regression analyses were performed for different levels of context and background noise due to the possibility that individual factors which influence performance in one condition may not necessarily influence performance in another condition, or may influence it to a lesser degree.

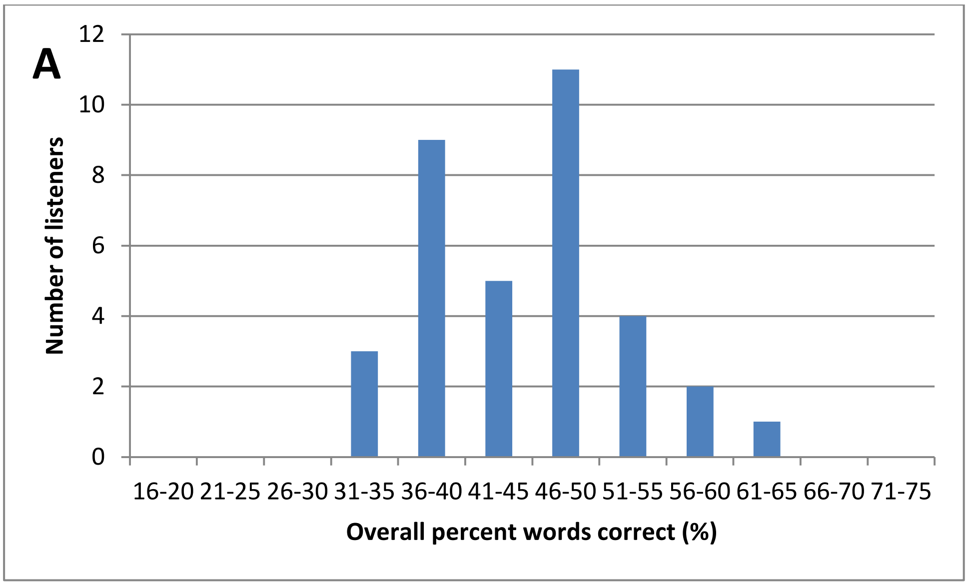
To determine if the influence of vocabulary on speech recognition scores differed according to the measure employed, the PPVT or the WAIS, the main analyses were repeated with the scores from each vocabulary test separately. A second reason that separate analyses were run for PPVT and WAIS vocabulary scores was the strong correlation between the two measures ($r = 0.64$, $p < 0.001$).

Finally, to ensure that variance contributions from age or working memory were not being obscured by vocabulary in the hierarchical regression analyses, further analyses were conducted with age and working memory respectively entered as the first predictor variable, and vocabulary entered third.

3. Results

3.1 Speech recognition

Variation was observed overall across both levels of predictability and in speech recognition performance in both HP and LP conditions in terms of the percentage of words accurately recognised. Figure 1 (A – C) displays the frequency distribution of speech recognition scores overall and for each predictability condition collapsed across all SNRs.



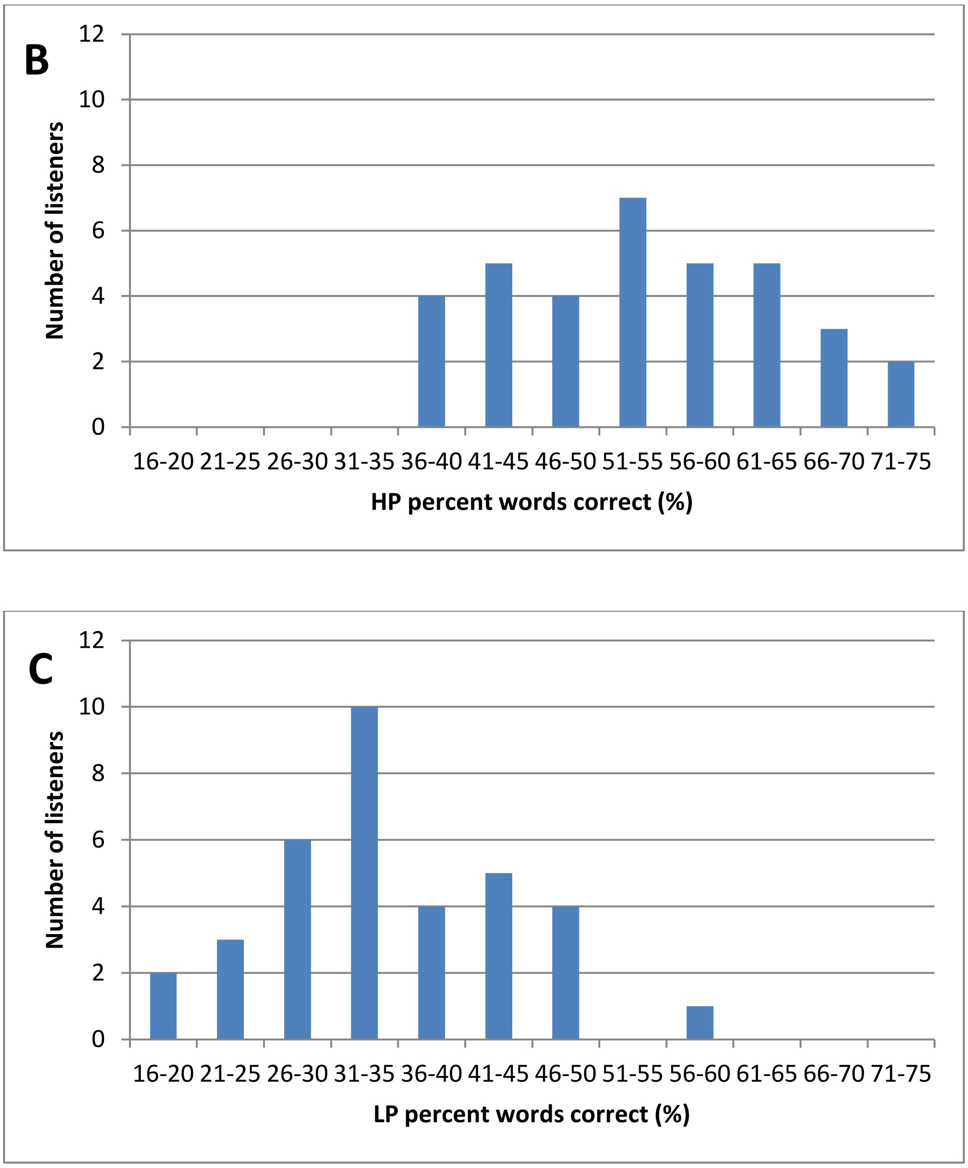


Figure 1: Histograms showing the frequency distribution of speech recognition scores in percent words accurately recognised: overall (A) and for the high predictability (B) and low predictability (C) conditions.

Overall, with results averaged across both levels of predictability and all SNRs, speech recognition performance ranged between 34% and 63.1% of the stimuli words correctly identified (mean accuracy of 45.8%; SD = 7.4). As expected, speech recognition accuracy was significantly higher for HP than LP stimuli t (34) = 11.44, p < 0.0001. LP speech recognition scores ranged from 16.9% to 57.5% (mean accuracy 35%; SD = 8.9), while HP speech recognition performance ranged from 38.3% to 73.7% (mean accuracy 54.4%; SD = 9.8).

The means and SDs of speech recognition performance in rationalised arcsine units (RAU) overall and for both LP and HP stimuli sets are displayed in Table 3 below. All further statistical analyses were performed using the RAU scores.

	Mean Score in RAU (SD)		
Signal-to-noise ratio	**LP**	**HP**	**OVERALL**
Overall Mean	36.2 (8.7)	54.5 (9.3)	46.1 (6.9)
4 dB	72.8 (14.9)	97.5 (7.9)	86.8 (9.4)
0 dB	45.7 (13.5)	74.1 (17.4)	61.8 (11.6)
-4 dB	17.9 (15.1)	37.8 (18.8)	30.4 (14.1)
-8 dB	-5.3 (10.8)	0.4 (14.9)	-1.9 (12.2)

Table 3: Mean scores with standard deviations in parenthesis for all participants in rationalized arcsine units (RAU) overall and for HP and LP conditions individually, across all SNRs and at each SNR individually.

3.2 Speech recognition performance in the high predictability versus low predictability conditions

Figure 2 compares mean listening recognition scores for HP and LP stimuli at each SNR. It demonstrates that listeners successfully recognised a greater proportion of the HP than LP stimuli at every SNR. This was confirmed statisitcally, with a two-way repeated measures ANOVA revealing a significant main effect of predictability condition on RAU listening recognition scores, $F(1, 34) = 119.14$, $p < 0.0001$.

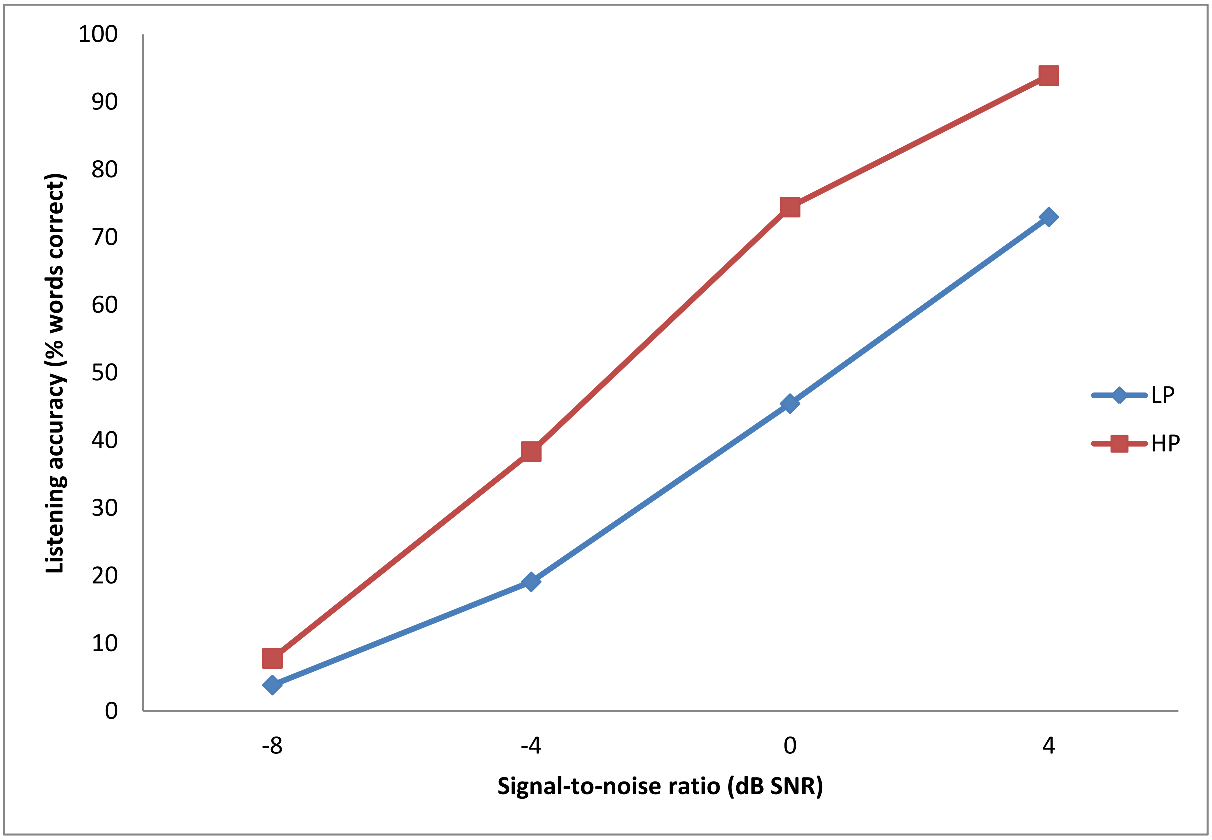


Figure 2: Line graph comparing mean speech recognition performance for all participants in the high predictability (HP) and low predictability (LP) conditions at each signal-to-noise (SNR) ratio in raw percent words accurately recognised.

Pairwise comparisons demonstrated that listening recognition performance was significantly better in the HP than the LP condition at each SNR except the poorest of -8 dB SNR (all pairwise tests $p < 0.0001$). Also in line with expectations, the mean proportion of words accurately recognised became successively greater as intelligibility improved across the four SNR conditions for both HP and LP stimuli. There was a significant main effect of SNR, $F(3, 102) = 687.96$, $p < 0.0001$. Pairwise comparisons confirmed that listeners were significantly more accurate at higher SNRs (all pairwise comparisons $p < 0.0001$) in both predictability conditions. Finally, a significant interaction existed between predictability and SNR, $F(3, 102) = 12.34$, $p < 0.0001$. This interaction reflects that there was no significant difference between LP and HP score in the -8 dB SNR condition, and as SNR improved HP scores increased more than LP scores.

3.3 Working memory and vocabulary tasks

The average scores for the vocabulary and working memory tasks were close to the normative means. The mean PPVT standard score for all listeners was 109.5 (SD = 10.5). WAIS scores were converted to age-adjusted scaled scores ranging from 0 - 19. The mean WAIS vocabulary score was 10.4 (SD = 1.3). The raw scores from the forward and backward digit span and digit sequencing tasks were also converted to age-adjusted scaled scores and then averaged to give a scaled working memory score in the range of 0-19. The mean working memory score was 10.6 (SD = 1.8).

3.4 Relationship between participant variables

The four participant variables which had been included in the study on the basis of prior evidence of being related to listening in adverse conditions were: vocabulary, both receptive as assessed by the PPVT (1) and productive as assessed by the WAIS (2), working memory (3) and age in chronological years (4). Table 4 displays inter-correlations between these predictor variables for the thirty-five participants. Receptive vocabulary as assessed by the PPVT-IV was moderately correlated with expressive vocabulary, as assessed by the WAIS vocabulary subtest ($p < 0.0001$). A modest negative correlation between working memory and age was also revealed ($p < 0.05$). Age was not significantly correlated with either expressive or receptive vocabulary. Correlations between working memory and receptive and expressive vocabulary did not reach significance.

Predictor variable	Receptive vocabulary	Expressive vocabulary	Working memory
Expressive vocabulary	0.64**		
Working memory	0.28	0.19	
Age	- 0.07	- 0.12	- 0.38*

*Table 4: R values of correlations between predictor variables (** = p < 0.0001; * = p < 0.05).*

3.5 Relationship between listening recognition performance and the predictor variables

This analysis began by investigating the simple pairwise relationships between the predictor variables and listening recognition performance. Moderate positive correlations were evident between PPVT vocabulary score and overall listening score (r = 0.59, p < 0.01), HP listening score (r = 0.47, p < 0.01), and LP listening score (r = 0.52, p < 0.01). WAIS vocabulary scores were also moderately correlated with overall listening recognition score (r = 0.49, p < 0.01), and LP listening score (r = 0.46, p < 0.05), while the relationship between WAIS score and HP listening score was weaker (r = 0.33, p = 0.05). Correlations between overall listening recognition performance and both working memory (r = 0.21) and age (r = 0.12) in chronological years did not reach statistical significance. Figures 3 –6 show scatter plots displaying the association with linear regression lines for overall RAU listening recognition scores collapsed across all conditions and PPVT vocabulary score, WAIS vocabulary score, working memory and age respectively.

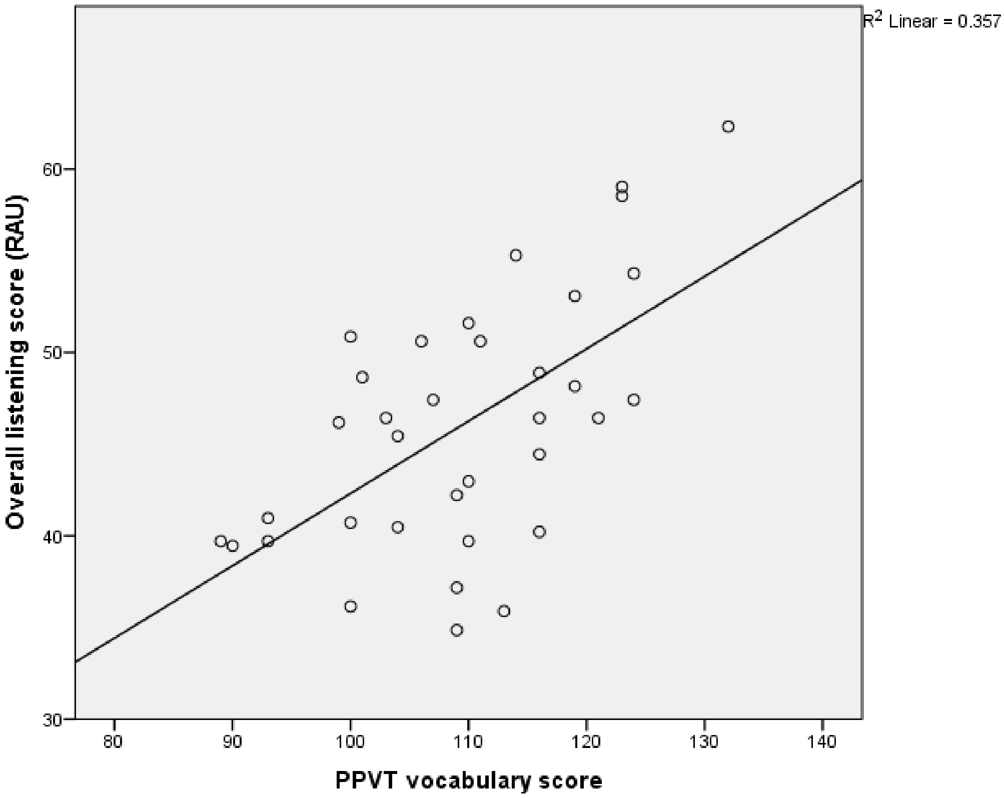


Figure 3: Scatterplot of PPVT vocabulary score and RAU speech recognition score overall with regression line.

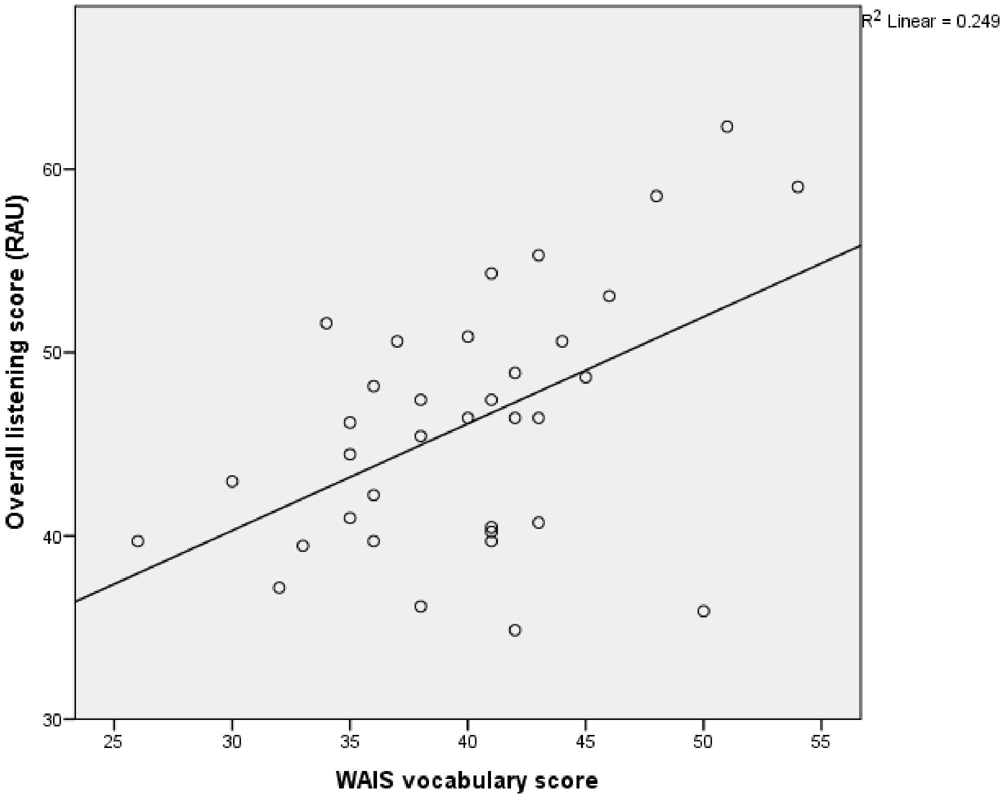


Figure 4: Scatterplot of WAIS vocabulary score and RAU speech recognition score overall for all participants with regression line.

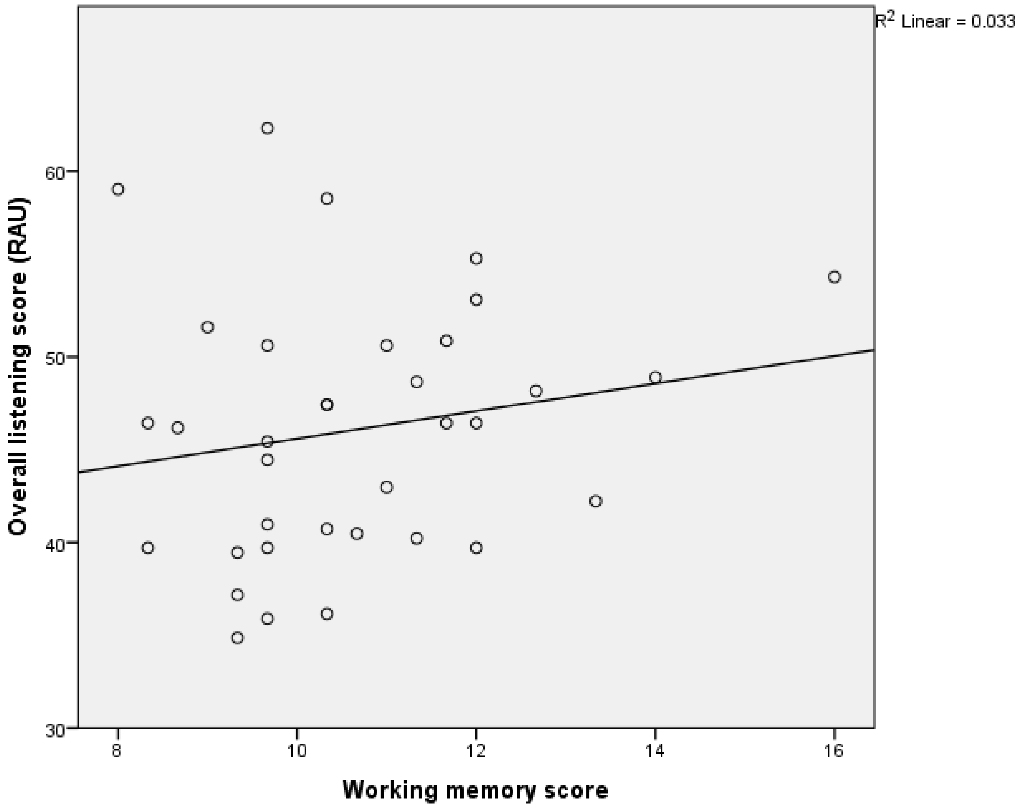


Figure 5: Scatterplot of working memory score and RAU speech recognition score overall for all participants with regression line.

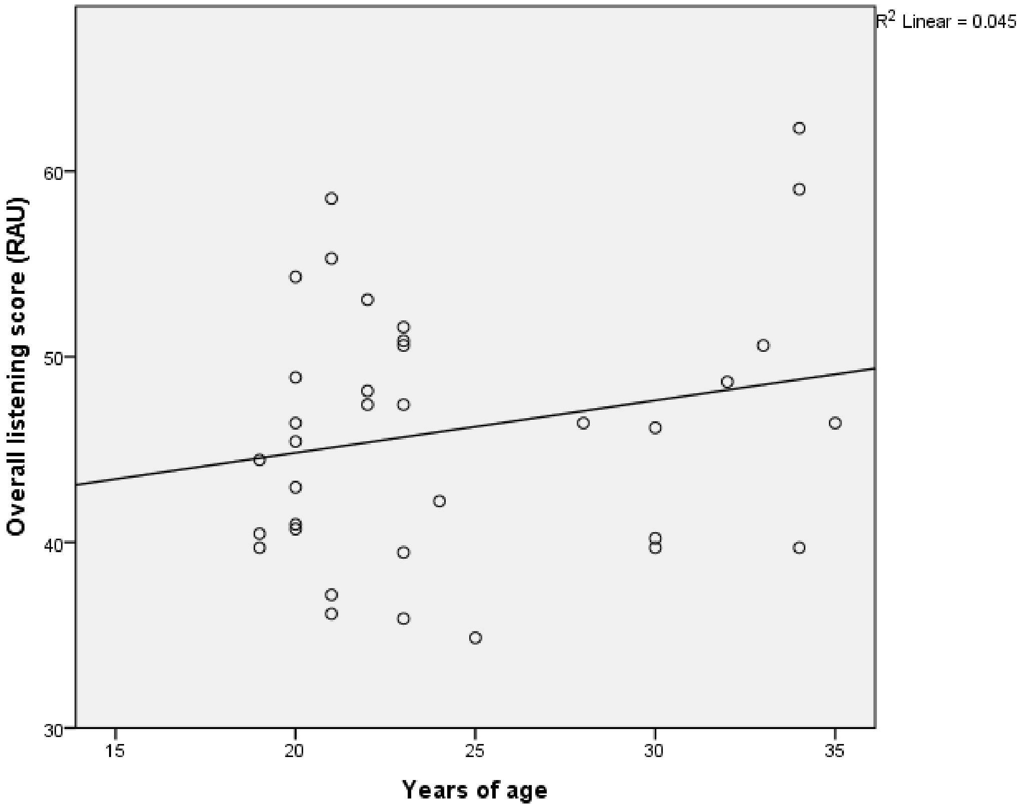


Figure 6: Scatterplot of age and RAU speech recognition score overall for all participants with regression line.

3.6 The effects of vocabulary, working memory and age on speech recognition

To investigate which listener characteristics could best explain variation in speech recognition accuracy, a series of hierarchical multiple regression analyses were performed.

For each hierarchical multiple regression analysis, speech recognition performance in rationalised arcsine units was the dependent variable. Predictor variables were entered into the model in the following order: (1) vocabulary (in standard scores for the PPVT or WAIS), (2) working memory, represented by the scaled score derived from the three tasks of the WAIS working memory subtest, (3) listeners' age in chronological years.

Tables 5-7 show the results of the regression analysis examining the relationship between the predictor variables and overall speech recognition score for both HP and LP stimuli across all SNRs, as well as separate analyses conducted with HP and LP stimuli speech recognition scores across all SNRs tested. The R^2 value representing the cumulative contribution to variability explained by each predictor variable is shown, as well as the change in R^2, which represents the further contribution made by each variable in addition to that already established. The level of statistical significance and the standardised regression coefficient (β) are also given.

The results of the regression analyses shown in Table 5 reveal that receptive vocabulary contributed significant variance to overall speech recognition score, and this effect remained when considering both HP and LP stimuli speech recognition scores separately. By contrast, neither age nor working memory contributed significant variance for overall speech recognition score or either level of stimuli predictability.

Predictability Condition	**Predictor variable**	***R^2***	**Change in *R^2***	***p***	***β***
Overall					
	PPVT Vocabulary	0.357	0.357	<0.0001	4.20
	Working memory	0.421	0.065	n.s.	1.87
	Age	0.424	0.003	n.s.	0.40
High predictability	*PPVT Vocabulary*	0.220	0.220	<0.005	0.48
	Working memory	0.247	0.027	n.s.	0.16
	Age	0.248	0.001	n.s.	-0.14
Low predictability	*PPVT Vocabulary*	0.273	0.273	<0.001	0.49
	Working memory	0.354	0.081	n.s.	0.35
	Age	0.375	0.021	n.s.	0.16

Table 5: Results of hierarchical multiple regression analyses overall and for HP and LP conditions separately with the predictor variables of PPVT vocabulary, working memory and age.

In order to investigate whether the role of the predictor variables differed according to the SNR, and therefore the degree of adverse conditions, separate regression analyses were conducted for the speech recognition results at each of the four SNRs employed for both LP and HP stimuli.

Table 6 presents the results of the regression analyses for the HP stimuli speech recognition results at each of the four SNRs applied during testing. The table shows that receptive vocabulary contributed significant variance to the speech recognition results at all SNRs except the highest SNR of +4 dB, which represents the best or least adverse condition in the study. Once receptive vocabulary had been entered into the model, working memory and age did not contribute significant variance to the speech recognition results at any SNR employed.

High Predictability Stimuli	**Predictor variable**	***R²***	**Change in *R²***	***p***	***β***
+4 dB SNR					
	PPVT Vocabulary	0.023	0.023	n.s.	0.04
	Working memory	0.067	0.044	n.s.	0.33
	Age	0.108	0.041	n.s.	0.22
0 dB SNR					
	PPVT Vocabulary	0.164	0.164	< 0.02	0.43
	Working memory	0.177	0.014	n.s.	-0.10
	Age	0.231	0.053	n.s.	0.25
-4 dB SNR					
	PPVT Vocabulary	0.165	0.165	< 0.02	0.44
	Working memory	0.170	0.005	n.s.	-0.11
	Age	0.174	0.004	n.s.	-0.07
-8 dB SNR					
	PPVT Vocabulary	0.158	0.158	< 0.02	0.34
	Working memory	0.169	0.011	n.s.	0.16
	Age	0.180	0.011	n.s.	0.12

Table 6: Results of hierarchical multiple regression analyses in the HP condition at each SNR with the predictor variables of PPVT vocabulary, working memory and age.

For the LP speech recognition results, Table 7 illustrates that receptive vocabulary contributed significant variance to the speech recognition results at all SNRs except the lowest SNR of -8 dB, which represents the poorest or most adverse condition in the study. After receptive vocabulary had been accounted for, working memory and age did not contribute significant variance at the poorest SNR employed of – 8 dB or the two best SNRs employed of +4 dB and 0 dB. At -4 dB, however, both working memory and age showed significant contributions to the speech recognition scores.

Low Predictability Stimuli	**Predictor variable**	***R^2***	**Change in *R^2***	***p***	***β***
+4 dB SNR					
	PPVT Vocabulary	0.183	0.183	< 0.01	0.39
	Working memory	0.187	0.004	n.s.	0.10
	Age	0.191	0.004	n.s.	0.07
0 dB SNR					
	PPVT Vocabulary	0.277	0.277	< 0.001	0.47
	Working memory	0.282	0.024	n.s.	0.18
	Age	0.325	0.038	n.s.	0.23
-4 dB SNR					
	PPVT Vocabulary	0.155	0.155	< 0.02	0.33
	Working memory	0.155	0.000	n.s.	0.30
	Age	0.518	0.363	<0.0001	0.66
-8 dB SNR					
	PPVT Vocabulary	0.000	0.000	n.s.	-0.03
	Working memory	0.010	0.010	n.s.	0.12
	Age	0.012	0.002	n.s.	0.04

Table 7: Results of hierarchical multiple regression analyses in the LP condition at each SNR with the predictor variables of PPVT vocabulary, working memory and age.

To summarise, receptive vocabulary contributed significant variance to the overall speech recognition scores and both HP and LP speech recognition scores. When SNR conditions were considered separately, vocabulary contributed significant variance to both HP and LP recognition scores at all SNRs utilised except HP stimuli at 4 dB SNR and LP stimuli at -8 dB SNR. The conditions in which vocabulary did not contribute variance, namely low context stimuli at a very poor SNR and high context stimuli at a relatively good SNR, represent the most adverse and the least adverse of all the conditions employed in the study.

The two remaining predictor variables entered into the modelling, working memory and age did not contribute unique variance to the overall speech recognition scores or either HP or LP speech recognition scores. Working memory and age also did not contribute significant variance when each SNR condition was considered in isolation, except for LP stimuli at -4 dB SNR.

The three chief analyses were repeated with productive vocabulary WAIS scores as the vocabulary variable (see Table 8 below). Productive vocabulary as assessed by the WAIS contributed significant variance to speech recognition scores overall and to speech recognition scores in the HP and LP conditions considered individually. The standardised coefficient values were lower than those achieved by the PPVT scores, however.

Predictability Condition	**Predictor variable**	***R²***	**Change in *R²***	***P***	***β***
Overall					
	WAIS Vocabulary	0.249	0.249	< 0.005	0.42
	Working memory	0.278	0.029	n.s.	0.27
	Age	0.314	0.036	n.s.	0.22
High predictability	*WAIS Vocabulary*	0.111	0.111	= 0.05	0.27
	Working memory	0.137	0.026	n.s.	0.23
	Age	0.156	0.019	n.s.	0.16
Low predictability	*WAIS Vocabulary*	0.216	0.216	< 0.01	0.37
	Working memory	0.242	0.026	n.s.	0.27
	Age	0.298	0.056	n.s.	0.28

Table 8: Results of hierarchical multiple regression analyses in the overall and for HP and LP conditions separately with the predictor variables of WAIS vocabulary, working memory and age.

To rule out the possibility that unique variance from working memory or age was being obscured by vocabulary in the analyses, further hierarchical multiple regression analyses were run with the variables entered in the following orders: (a) age, working memory, receptive vocabulary; (b) working memory, age, receptive vocabulary. When age was entered as the first predictor variable in the analysis, it did not contribute significant variance to overall speech recognition scores (R^2 = 0.04, n.s.), HP scores (R^2 = 0.02, n.s.) or LP scores (R^2 = 0.06, n.s.). Likewise, when working memory was entered as the first predictor variable, the measure did not contribute significant variance to overall speech recognition scores (R^2 = 0.03, n.s.), HP scores (R^2 = 0.02, n.s.) or LP scores (R^2 = 0.03, n.s.).

4. Discussion

The purpose of this study was to investigate the influence of receptive and productive vocabulary, working memory and age on speech recognition performance in multitalker babble in younger adults with normal hearing acuity. The two main aims of the study were: (1) to confirm the association between vocabulary knowledge and speech recognition in adverse conditions, (2) to examine whether this association is due to vocabulary knowledge being related to top-down linguistic processing of degraded speech.

It was expected that speech recognition accuracy would be lower for LP than HP stimuli and would deteriorate as SNR became successively poorer. Furthermore, it was hypothesised that listeners with larger receptive vocabularies would be more accurate at recognising both HP and LP stimuli, but that vocabulary would be more strongly related to speech recognition accuracy in the HP condition. It was also anticipated that the influence of receptive vocabulary would vary according to the degree of adverse listening conditions. We hypothesised that when listening conditions were relatively favourable or very unfavourable receptive vocabulary would not influence speech recognition accuracy. At intermediate levels of adverse listening conditions, however, we expected receptive vocabulary to exert an influence on speech recognition accuracy.

By contrast, it was hypothesised, based on prior studies, that productive vocabulary knowledge (WAIS vocabulary score) would not be related to speech recognition in noise. Due to the brevity of the target phrases employed, it was hypothesised that working memory would not be related to speech recognition accuracy.

The primary findings of the study indicated that: (1) listeners with greater receptive vocabulary were able to accurately recognise more speech in both the HP and LP conditions, (2) vocabulary was not more strongly associated with speech recognition performance in the HP than the LP condition, (3) receptive vocabulary knowledge did not influence speech recognition accuracy in the least favourable condition or the most favourable condition, but (4) greater receptive vocabulary knowledge was related to superior speech recognition accuracy at all intermediate levels of listening favourability, (5) productive vocabulary (WAIS IV) vocabulary was positively related to speech recognition in both the HP and LP conditions, (6) working memory did not influence speech recognition accuracy in either the HP or LP condition, (7) age did not influence speech recognition accuracy in either the HP or LP condition overall, but was positively associated with speech recognition accuracy in the LP condition at -4 dB SNR.

Secondary findings of the current study were that listeners were more accurate at recognising HP than LP speech and that speech recognition accuracy decreased as SNR deteriorated.

Each of the findings will be discussed in the following section.

4.1 Stimulus predictability and signal-to-noise ratio

The finding that listeners were more accurate with HP than LP stimuli was expected and in line with the existing literature (e.g. Hutchinson, 1989; McAuliffe et al, 2011). The listeners were able to exploit the semantic and contextual cues to aid their recognition of the speech in the HP condition, but such cues were absent in the LP condition.

The finding that speech recognition accuracy improved as SNR was incrementally increased is consistent with extensive psychoacoustic literature (Miller, 1947). As the level of the speech signal relative to the background speech is increased, it becomes less demanding for listeners to segregate the target speech.

4.2 The influence of vocabulary knowledge on speech recognition accuracy

The finding that receptive vocabulary influences listening proficiency in adverse conditions supports previous research which has also found this effect (McAuliife et al; 2013; Benard et al, 2014; Tamati et al, 2013; Janse & Adanks, 2012).The current study shows the effect of receptive vocabulary on listening in adverse conditions exists for both high and low predictability stimuli, and is evident in adverse conditions associated with masking by multitalker babble.

Two things arising from the current study are important to note. Firstly, it is unlikely that listeners with larger vocabularies were more successful on the speech recognition task because they were more familiar with or knew more of the words in the target sentences. As mentioned, 100% of the HP and 97.1% of the LP words fell within the 6000 most common words in English according to the COCA and BNC corpora. Zechmeister, Chronis, Cull, D'Anna and Healy (1995) calculated that the receptive vocabulary size of a university graduate is approximately 17,000 word families, while the receptive vocabulary of a first year university student is approximately 12,000 word families. Similarly, Goulden, Nation, and Read (1990) found that the receptive vocabulary size of university-educated speakers ranged between 13,200 and 20,700 base words, with an average of 17,200 base words.

Of the participants in the current study, 24 had completed university education, while the remainder were current undergraduate students. Therefore, it is assumed the participants' receptive vocabularies were around 17,000 to 20,000 words, but possibly as low as 12,000. Given this, it is unlikely any participant would not be familiar with the 6,000 most common words in English, which include the vast majority of the target stimuli words.

Hence, it appears that vocabulary knowledge has a general effect on speech recognition which is not simply due to familiarity with the target words. Secondly, the effect of vocabulary was consistent for HP and LP stimuli, implying that listeners with larger vocabularies were not only more accurate because they were better able to take advantage of contextual and semantic predictability.

4.3 Reasons for the relationship between vocabulary knowledge and speech recognition in adverse conditions

If, we assume, the target words fell within the receptive vocabularies of all the participants, why were participants with larger vocabularies more accurate at recognising the words than participants with smaller vocabularies?

One possibility is that listeners with greater vocabulary knowledge are able to deploy superior top-down processing resources to accurately separate and/or reconstruct a degraded speech signal. McAuliffe et al (2013) speculated that this heightened top-down linguistic processing of degraded speech could be a result of listeners with greater vocabulary knowledge having accumulated greater experience and familiarity with language. McAuliffe et al put forward that listeners with larger vocabularies are better able to exploit intelligible fragments and redundancies in the of the speech stream to draw accurate lexical hypotheses and reconstruct the signal, relating this to the 'glimpsing' theory of speech segregation in noise (Cooke, 2005).

In addition to exploiting intelligible glimpses, listeners with superior vocabulary knowledge may also be better able to take advantage of the acoustic, phonetic and linguistic redundancies in the speech signal, which would afford them greater listening success (Assmann & Summerfield, 2003).

4.4 The differential influence of receptive vocabulary on speech recognition accuracy according to degree of listening difficulty

Support for the notion that listeners with larger vocabularies have superior top-down linguistic processing is provided by a finding which is unique to the current study: the influence of receptive vocabulary knowledge on speech recognition accuracy in noise varies according to the degree of adverse conditions. We found that, in line with our hypotheses, receptive vocabulary did not exert influence of speech recognition accuracy when listening conditions were favourable or very unfavourable, but did exert significant influence at intermediate levels of listening difficulty.

We speculate that the reason vocabulary did not influence speech recognition in relatively favourable listening conditions is that there was a comparatively small amount of variability between listeners in terms of speech recognition accuracy. In favourable listening conditions, there is typically much less between individual variability than when listening conditions are adverse because listening is effortless and top-down influence on processing is slight (Gilbert et al, 2013; Tamati et al, 2013; Assman & Summerfield, 2004).

The most favourable listening condition in the current study, in which no influence of vocabulary was present, included sentences with high semantic and syntactic predictability at a SNR ratio of 4 dB HL. In these listening conditions a speech recognition score of close to 100% would be expected (McArdle & Hnath-Chisolm, 2009). This was borne out in the current study, with a mean raw speech recognition accuracy score in the 4 dB SNR HP condition of 94%. Importantly, the variability in speech recognition scores was also low, with the lowest RAU standard deviation (7.9) of all eight conditions. Therefore, we conjecture that the reason vocabulary did not influence speech recognition accuracy in this condition was that the participants did not find this condition challenging and performed with minimal variability.

The most unfavourable condition, which also did not show an influence of vocabulary on performance, included sentences with low semantic and syntactic predictability at an SNR ratio of -8 dB SNR. In these listening conditions, speech recognition accuracy is likely to be very poor and our results confirmed this with a mean raw speech recognition score of 3.9%.

A potential reason for the lack of influence of vocabulary on speech recognition accuracy in the LP -8 dB SNR condition is that, similarly to the HP 4 dB SNR condition, there was a lack of variability in performance between listeners. However, while the variability in the LP -8 dB SNR condition was the second lowest of the eight conditions, the RAU SD (10.9) was not substantially lower than the average RAU SD of the other conditions (13.2).

Another possible explanation for the lack of influence of vocabulary in the LP -8 dB SNR condition is based on the hierarchical framework of speech segmentation cues proposed by Mattys et al (2005), who investigated the amount of weight listeners apply to lexical, sub-lexical and prosodic cues in segmenting the speech stream. According to their model, when listening conditions are good listeners rely predominately on knowledge-derived lexical cues. However, when conditions deteriorate and lexical information becomes ambiguous, listeners resort to sub-lexical cues to inform segmentation. When listening conditions worsen further, Mattys et al. (2005) demonstrated that listeners may rely on syllabic stress to determine word boundaries. Considering the listening conditions applied in the current study in the context of the model provided by Mattys et al (2005), it is possible that in the -8 dB LP condition, both lexical and sub-lexical cues have become sufficiently degraded that all listeners regardless of their vocabulary knowledge rely on syllabic stress to parse the speech stream, meaning that no advantage of superior top-down processing would exist.

Further research in our laboratory is planned to determine whether the lack of influence of vocabulary knowledge on speech recognition in extremely adverse listening conditions is due to lack of between listener variability or a result of listeners changing their listening strategies as the signal deteriorates. To achieve this, the speech recognition errors made by listeners will be analysed to determine what kind of listening strategy was being employed.

As well as introducing degrees of adverse conditions, the current study was unique in directly comparing the influence of vocabulary with HP and LP stimuli. If top-down linguistic processing skill is the reason vocabulary knowledge is associated with speech recognition proficiency, we might expect listeners with larger vocabularies to outperform those with smaller vocabularies even more in the HP condition than in the LP condition. The results did not support this hypothesis, however. In fact, vocabulary accounted for marginally more variance in speech recognition scores in the LP condition than the HP condition. Despite this, the result does not rule out top-down linguistic processing as an underlying source of the relationship between vocabulary and speech perception, because speech contains many layers of redundancy. In addition to those provided by semantic and contextual predictability, there are also acoustic, phonetic and phonological redundancy cues (Assmann & Summerfield, 2004). The superior exploitation of redundancies and intelligible glimpses performed by listeners with larger vocabularies could involve, for example, better mapping of sound sequences onto words or better use of the amplitude modulated envelope of speech (Assmann & Summerfield, 2004).

4.5 The influence of working memory and age on speech recognition accuracy

The current study found that working memory as assessed by a WAIS composite score did not influence speech recognition accuracy in any condition, confirming our hypothesis. This result is consistent with McAuliffe et al (2013), who used the same LP stimuli as the current study, but conflicts with other research (Tamati et al, 2013; Janse & Adank, 2012) found that working memory had an influence on speech recognition in adverse conditions.

McAuliffe et al (2013) attributed the lack of an influence of working memory on speech recognition performance to the brevity of the stimuli. The stimuli in the current study and that of McAuliffe et al were all six syllables in length, ranging from three to six words per phrase, so it is possible that the use of short sentences precluded an effect of working memory on listening proficiency. Further studies using longer sentences or sentences with a range of lengths are required to determine whether any influence of working memory varies according to stimulus length.

Another possibility for a lack of influence of working memory on speech recognition performance in the current study is a lack of variability in the working memory scores. Raw WAIS subtest scores are converted to scaled scores with a mean of 10 and a standard deviation of 3. The mean working memory score for the participants in the current study was 10.6 with a standard deviation of 1.8, showing the participants performed slightly above average and had less variance in their scores compared to the normative population.

Studies which have succeeded in finding an influence of working memory on speech recognition in noise have tended to include older participants who may have greater variability in working memory capacity (Akeroyd, 2008), and have also used measures of working memory other than the WAIS digit span and digit sequencing subtests. The measures of working memory which have given significant results are the reading span test (Foo, Rudner, Ronnberg & Lunner, 2007; Rudner, Foo, Ronnberg & Lunner, 2007) and letter monitoring tasks (Gatehouse, Naylor & Elberling, 2003; Foo et al, 2007). Therefore, it is possible that the finding of the current study with regard to working memory is associated with the measure used to assess working memory and/or the inclusion of only younger participants.

Age did not exert a significant influence on overall LP and HP speech recognition scores in the current study. Only listeners aged 18 to 35 with normal hearing acuity, were employed, so the result that age did not influence speech recognition performance was expected. Indeed, even some studies comparing speech recognition in adverse conditions in younger and older listeners with normal hearing acuity have failed to find an effect of age on speech recognition accuracy in adverse condition (McAuliffe et al, 2011; 2013).

In the -4dB SNR LP condition, age did exert a significant influence, however. The result that age influenced speech recognition accuracy in one of the 8 SNR and predictability conditions was unexpected and the reason for it is not immediately apparent. The data was inspected for outliers in the -4 dB SNR LP condition. Three cases were identified as possibly exerting a significant influence on the regression model. However, when these cases were removed and the regression was repeated, age still accounted for significant variance in speech recognition scores.

As mentioned, the participants were exposed to the various SNRs in a random order. It might be expected that performance would improve as more trials were completed, so one possibility is that older participants tended to hear the -4 dB SNR LP condition stimuli later than younger participants, and benefited from a learning effect.

4.6 Limitations of the study and avenues for further research

The findings of the current study should be considered within the context of its limitations.

Firstly, the participants in the current study performed somewhat better than average and with less variability compared to the norm-referenced population for all the cognitive measures including receptive vocabulary, productive vocabulary, and working memory. Further research is required which includes participants whose average level of vocabulary knowledge more closely resembles the norm and has a greater degree of variability.

In order to achieve this, future studies will likely have to overcome the second shortcoming of the current study, which is a small sample size. Further research would benefit from recruiting a larger group of participants with a more equal division of males and females.

A third limitation of the current study is that a limited number of cognitive assessments were included. Apart from vocabulary, working memory was the only cognitive predictor assessed due to its tendency to influence speech recognition in prior research. Future research could benefit from having a wider range of cognitive measures to preclude the possibility that vocabulary is highly inter-correlated with other cognitive measures or general intelligence.

A further limitation of the current study is that the listening task was not natural, because the listeners were required to rely on their sense of hearing alone to recognise the speech. Outside a laboratory setting, most speech recognition occurs when listeners have are able to see the speaker. Further research could aim to determine whether vocabulary still exerts an influence on speech recognition accuracy with an audiovisual signal.

The finding of the current study that vocabulary knowledge exerts influence speech on recognition in moderately adverse listening conditions, but not in favourable or very adverse conditions, poses questions which require further investigation. One strategy for further research could be to examine lexical segmentation.

Lexical segmentation is the process of segmenting the speech signal into individual words (Jusczyk & Luce, 2002), a key process in speech recognition in which lexical knowledge is believed to be important (McAuliffe et al, 2013). To examine a listener's lexical segmentation process, one strategy is analyse their lexical boundary errors (LBEs). An LBE is "an incorrect deletion or insertion of a word or lexical boundary" (McAuliffe et al, 2013, p. 1363). LBEs can be analysed to determine whether a listener based their lexical segmentation on syllabic stress cues or lexical cues. Potential avenues for future research could include: (1) determining whether the lexical segmentation strategies employed by listeners with superior vocabulary knowledge differ from those with more limited vocabulary knowledge, (2) determining whether the lexical segmentation strategy employed by listeners differs according to an interaction between the degree of adverse conditions experienced and their level of vocabulary knowledge.

It might be hypothesised that listeners with greater vocabulary knowledge are less likely to rely on syllabic stress cues and more often employ a lexically-based segmentation strategy than listeners with more limited vocabulary knowledge. Another hypothesis could be that listeners with greater vocabulary knowledge are more likely than those with more limited vocabularies to use a lexically based segmentation strategy when listening conditions are moderately adverse, but not when conditions are favourable (when both use a lexically based strategy) or very adverse (when both rely on sublexical cues).

The current study employed only phrasal stimuli which contained words with high lexical frequency. A further opportunity for further research could be to determine whether the influence of vocabulary knowledge on speech recognition differs according to word difficulty by including stimuli with varying levels of lexical frequency.

4.7 Conclusion

This study examined the influence of receptive and productive vocabulary knowledge, working memory and age on speech recognition in noise in younger listeners with normal hearing acuity. Performance was determined on the basis of percent words accurately recognised with both HP and LP phrases at various SNRs.

As hypothesised, receptive vocabulary knowledge influenced speech recognition performance in both HP and LP conditions. Also in line with our hypotheses, receptive vocabulary knowledge did not influence speech recognition performance when conditions were relatively favourable or very unfavourable, but did exert significant influence at intermediate levels of adverse conditions. Unexpectedly, productive vocabulary also exerted significant influence on HP and LP speech recognition performance.

The results suggest that listeners with greater vocabulary knowledge are more proficient listeners in adverse listening conditions. We speculate that the reason for this superiority heightened top-down processing of the degraded speech.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, *47*(S2), S53-S71.

Alamsaputra, D. M., Kohnert, K. J., Munson, B., & Reichle, J. (2006). Synthesized speech intelligibility among native speakers and non-native speakers of English. *Augmentative and Alternative Communication*, *22*(4), 258-268.

Altieri, N., Gruenenfelder, T., & Pisoni, D. B. (2010). Clustering coefficients of lexical neighborhoods: Does neighborhood structure matter in spoken word recognition? *The Mental Lexicon*, *5*(1), 1-21.

Assmann, P., & Summerfield, Q. (2004). The perception of speech under adverse conditions. In *Speech processing in the auditory system* (pp. 231-308). New York: Springer.

Başkent, D. (2010). Phonemic restoration in sensorineural hearing loss does not depend on baseline speech perception scores. *The Journal of the Acoustical Society of America*, *128*(4), EL169-EL174.

Beer, J., Kronenberger, W. G., & Pisoni, D. B. (2011). Executive function in everyday life: implications for young cochlear implant users. *Cochlear Implants International*, *12*(s1), S89-S91.

Benard, M. R., Mensink, J. S., & Başkent, D. (2014). Individual differences in top-down restoration of interrupted speech: Links to linguistic and cognitive abilities. *The Journal of the Acoustical Society of America*, *135*(2), EL88-EL94.

Benichov, J., Cox, L. C., Tun, P. A., & Wingfield, A. (2012). Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability and cognitive function. *Ear and Hearing*, *32*(2), 250-256.

Borrie, S. A., McAuliffe, M. J., & Liss, J. M. (2012). Perceptual learning of dysarthric speech: A review of experimental studies. *Journal of Speech, Language, and Hearing Research*, *55*(1), 290-305.

Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *The Journal of the Acoustical Society of America*, *106*(4), 2074-2085.

Brungart, D. S., Chang, P. S., Simpson, B. D., & Wang, D. (2009). Multitalker speech perception with ideal time-frequency segregation: Effects of voice characteristics and number of talkers. *The Journal of the Acoustical Society of America*, *125*(6), 4006-4022.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: the influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, *96*(4), 671-681.

Carhart, R., & Tillman, T. W. (1970). Interaction of competing speech signals with hearing losses. *Archives of Otolaryngology,91*, 274–279.

Cooke, M. (2003). Glimpsing speech. *Journal of Phonetics*, *31*(3), 579-584.

Cooke, M. (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*, *119*(3), 1562-1573.

Cooke, M., Barker, J., Cunningham, S., & Shao, X. (2006). An audio-visual corpus for speech perception and automatic speech recognition. *The Journal of the Acoustical Society of America*, *120*(5), 2421-2424.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, *31*(2), 218-236.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech, Language, and Hearing Research*, *12*(2), 246-269.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hearing Research*, *229*(1), 132-147.

Dillon, H. (2012). *Hearing Aids*. New York: Thieme.

Dubno, J. R., Ahlstrom, J. B., & Horwitz, A. R. (1999). Use of context by young and aged adults with normal hearing. *The Journal of the Acoustical Society of America*, *107*(1), 538-546.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test: PPVT-4B*. NCS Pearson.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research*, *47*(2), 421-436.

Festen, J. M., & Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing. *The Journal of the Acoustical Society of America*, *88*(4), 1725-1736.

Foo, C., Rudner, M., Rönnberg, J., & Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology*, *18*, 553-566.

Gatehouse, S., Naylor, G., & Elberling, C. (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *International Journal of Audiology*, *42*, S77-S85.

Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013). Development, Reliability and Validity of PRESTO: A New High-Variability Sentence Recognition Test. *Journal of the American Academy of Audiology*, *24*(1), 26-36.

Gordon-Salant, S., & Fitzgibbons, P. J. (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *Journal of Speech, Language and Hearing Research*, *40*(2), 423-431.

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, *11*(4), 341-363.

Heaps, H. S. (1978). *Information retrieval: Computational and theoretical aspects*. Academic Press, Inc..

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE [Computer software]. *Wellington, New Zealand: Victoria University of Wellington*.

Howes, D. (1957) On the relationship between the intelligibility and the frequency of occurrence of English words. *Journal of the Acoustical Society of America, 29,* 296–305.

Humes, L. E. (1996). Speech understanding in the elderly. *Journal of the American Academy of Audiology*, *7*, 161-167.

Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *The Journal of the Acoustical Society of America*, *112*(3), 1112-1132.

Hutchinson, K. M. (1989). Influence of sentence context on speech perception in young and older adults. *Journal of Gerontology*, *44*(2), P36-P44.

Janse, E., & Adank, P. (2012). Predicting foreign-accent adaptation in older adults. *The Quarterly Journal of Experimental Psychology*, *65*(8), 1563-1585.

Jerger, J., Jerger, S., & Pirozzolo, F. (1991). Correlational analysis of speech audiometric scores, hearing loss, age, and cognitive abilities in the elderly. *Ear and Hearing*, *12*(2), 103-109.

Johnson, C. D., & Seaton, J. B. (2011). *Educational audiology handbook*. Cengage Learning.

Jusczyk, P. W., & Luce, P. A. (2002). Speech perception and spoken word recognition: Past and present. *Ear and Hearing*, *23*(1), 2-40.

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*, *61*(5), 1337-1351.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, *54*(3), 399-436.

Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *The Journal of the Acoustical Society of America*, *104*(4), 2457-2466.

Luce, P. A., & McLennan, C. T. (2005). Spoken word recognition: The challenge of variation. In D. B. Pisoni & R. B. Remez (Eds.), *The Handbook of Speech Perception*, (pp. 590-609). Malden, MA: Blackwell.

Mason, H. M. (1946). Understandability of speech in noise as affected by region of origin of speaker and listener. *Communications Monographs*, *13*(2), 54-58.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953-978.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: a hierarchical framework. *Journal of Experimental Psychology: General*, *134*(4), 477.

McArdle, R., & Hnath-Chisolm, T. (2009). Speech audiometry. In J. Katz (Ed.), *Handbook of clinical audiology* (pp.64–79). Baltimore: Williams and Wilkins.

McAuliffe, M. J., Gibson, E. M., Kerr, S. E., Anderson, T., & LaShell, P. J. (2013). Vocabulary influences older and younger listeners' processing of dysarthric speech. *The Journal of the Acoustical Society of America*, *134*(2), 1358-1368.

McAuliffe, M. J., Wilding, P. J., Rickard, N. A., & O'Beirne, G. A. (2012). Effect of speaker age on speech recognition and perceived listening effort in older adults with hearing loss. *Journal of Speech, Language, and Hearing Research*, *55*(3), 838-847.

Miller, G. A. (1947). The masking of speech. *Psychological Bulletin*, *44*(2), 105-129.

Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America, 22,*167-173.

Munson, B. (2001). Relationships between vocabulary size and spoken word recognition in children aged 3 to 7. *Contemporary Issues in Communication Science and Disorders*, *28*, 20-29.

Nation, P. & Heatley, A. (1994). Range. *Victoria University of Wellington, NZ.[Computer program, available at http://www. vuw. ac. nz/lals/.]*.

Olsen, W. O., & Carhart, R. T. (1967). Development of test procedures for evaluation of binaural hearing aids. *Bulletin of Prosthetics Research,10*, 22-49.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America*, *97*(1), 593-608.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology*, *42*(S2), 11-16.

Pollack, I., Rubenstein, H., & Decker, L. (1959). Analysis of incorrect responses to an unknown message set. *The Journal of the Acoustical Society of America*, *32*(4), 454-457.

Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science*, *212*(4497), 947-949.

Rudner, M., Foo, C., Rönnberg, J., & Lunner, T. (2007). Phonological mismatch makes aided speech recognition in noise cognitively taxing. *Ear and hearing*, *28*(6), 879-892.

Schneider, B. A., Li, L., & Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology*, *18*(7), 559-572.

Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K. (2002). Listening in aging adults: from discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology*, *56*(3), 139.

Schlauch, R. S., & Nelson, P. (2009). Puretone evaluation. In J. Katz (Ed.), *Handbook of clinical audiology* (pp. 30-49). Baltimore: Lippincott Williams and Wilkins.

Sheldon, S., Pichora-Fuller, M. K., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-vocoded speech by younger and older adults. *The Journal of the Acoustical Society of America*, *123*(1), 489-499.

Sherbecoe, R. L., & Studebaker, G. A. (2004). Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units. *International Journal of Audiology*, *43*(8), 442-448.

Smaldino, J., Crandell, C., Kreisman, B., John, A., & Kreisman, N. (2009). Room acoustics and auditory rehabilitation technology. In J. Katz (Ed.) *Handbook of Clinical Audiology* (pp.745-775). Baltimore: Lippincott Williams and Wilkins.

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *The Journal of Neuroscience*, *32*(25), 8443-8453.

Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, *31*(4), 577-607.

Studebaker, G. A. (1985). A" rationalized" arcsine transform. *Journal of Speech and Hearing Research*, *28*(3), 455-462.

Tamati, T. N., Gilbert, J. L., & Pisoni, D. B. (2013). Some factors underlying individual differences in speech recognition on PRESTO: A first report. *Journal of the American Academy of Audiology*, *24*(7), 616-634.

Valente, M. I., & Valente, M. A. (2009). Hearing aid fitting for adults: selection, fitting, verification, validation. In J. Katz (Ed.) *Handbook of clinical audiology* (pp. 846-849). Baltimore: LippincWilliams & Wilkins.

Warren, R. M. (1984). Perceptual restoration of obliterated sounds. *Psychological bulletin*, *96*(2), 371.

Wechsler, D. (2008). Wechsler Adult Intelligence Scale - Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson*.

Wightman, F. L., Kistler, D. J., & O'Bryan, A. (2010). Individual differences and age effects in a dichotic informational masking paradigma). *The Journal of the Acoustical Society of America*, *128*(1), 270-279.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *The Journal of Neuroscience*, *32*(40), 14010-14021.

Zechmeister, E. B., Chronis, A. M., Cull, W. L., D'Anna, C. A., & Healy, N. A. (1995). Growth of a functionally important lexicon. *Journal of Literacy Research*, *27*(2), 201-212.

Appendix 1: Participant consent form

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

CONSENT FORM

Project Name: "Does vocabulary knowledge influence speech recognition in older listeners?"

Principal Investigator:

Joseph Dalrymple-Alford, Masters of Audiology Student

Associate Investigators:

Dr Megan McAuliffe, Associate Professor, Department of Communication Disorders

Dr Don Sinex, Senior Lecturer, Department of Communication Disorders

- I have read and I understand the information sheet dated 1st May 2013 for volunteers taking part in the study.

- I have had the opportunity to discuss this study with the researcher/s. I am satisfied with the answers I have been given.

- I understand that my participation in this study is confidential and that no material that could identify me will be used in any reports on this study.

- I have had time to consider whether to take part.

- I understand that taking part in this study is voluntary (my choice) and that I may withdraw from the study at any time. I am also aware that this will in no way affect my future interactions with the Department of Communication Disorders.

- This proposal has been reviewed and approved by the Department of Communication Disorders, University of Canterbury, and the University of Canterbury Human Ethics Committee Low Risk process.

I'm happy to be contacted for future studies	**YES/NO**
I consent to my speech recordings being stored and used in future studies that have received ethical clearance from the UC HEC and/or NZ Health and Disabilities Commission	**YES/NO**
I consent to the results of these assessments being made available for future studies if required	**YES/NO**
I give permission to the research team to access my previous audiological clinical records from the University of Canterbury or, if from another clinic, please put the clinic name and address here:	**YES/NO**
I wish to receive a copy of the results	**YES/NO**

I hereby consent to take part in the study:

Name (please print): ______________________________

Signature: ______________________ Date: ______________________

Project Explained By: ______________________________

Project Role: ______________________________

Signature ______________________ Date: ______________________

Appendix 2: Listener information sheet

INFORMATION SHEET – LISTENER GROUP

1.04.13

<u>Project Name:</u>

You are invited to take part in the research project titled: “Does vocabulary knowledge influence speech recognition?”

Please take the time to read this information sheet thoroughly and consider whether you would like to participate. Your participation is entirely voluntary (your choice). The following research team is conducting this study:

<u>Principal Investigator</u>:

Joseph Dalrymple-Alford, Master of Audiology Student

<u>Associate Investigators</u>:

Dr Megan McAuliffe, Associate Professor, Department of Communication Disorders

Dr Don Sinex, Senior Lecturer, Department of Communication Disorders

We are interested in how older people understand speech in noise and how this relates to vocabulary knowledge and memory. An understanding of how people understand speech in noise will be helpful for audiologists in the development of assessment and treatment plans.

Your ears will be examined, then earphones will be placed in your ears through which you will hear some beeps. The hearing assessment will be administered by a Masters of Audiology student in the soundproof testing booth at the Department of Communication Disorders.

Once the hearing test has been completed, you will be asked to undertake the following tasks:

(1) *Listening task:* Earphones or headphones will be fitted onto your head. You will hear sentences spoken by a man. Some of these sentences will be in quiet and some of them will be in noise. We will simply ask you to repeat back what you think you heard.

(2). *Cognitive tasks:* You will be asked to listen to a sequence of random digits and then repeating back the digits in the same you hear them or in reverse order and in order from lowest to highest.

(3) *Vocabulary tasks:* you will be asked to provide definitions for some words. You will also be asked to select by pointing to pictures words spoken by the tester.

In total, your involvement in this study will take approximately 60 minutes. All of the tests may be completed in one session, or over two sessions if you would prefer.

CONFIDENTIALITY

- Your privacy and confidentiality will be maintained at all times.
- All information will be kept in a locked filing cabinet at the Department of Communication Disorders, University of Canterbury. Only the researchers or research assistants involved in the project will have access to this information.
- The results of this project will be published; however, no material which could personally identify you will be used in any reports on this study.
- Feedback on individual assessment results will be provided at the time of testing.

- If you wish you will be advised of the results of the study.
- The research team may need to access your previous audiological clinical records, for which your consent will be required.

COMPENSATION FOR PARTICIPATION

You will receive a $20 petrol voucher for your participation in this study.

RISKS OF PARTICIPATION

There are no physical risks to participating in this project. Due to the length of the sessions you will be given as many breaks as you feel necessary. If you feel uncomfortable or unable to continue at any time you can withdraw from the study.

LOCATION

The hearing assessment will be conducted at the Department of Communication Disorders Speech and Hearing Clinic.

WITHDRAWING FROM THE STUDY

It is important to note that this study is voluntary and that you can withdraw from it at any time. This will in no way jeopardise any of your future dealings with the Department of Communication Disorders. If you choose to withdraw from the study, any data collected prior to withdrawal will not be used for research purposes without your consent.

ETHICS

This study has received ethical approval from the Human Ethics Committee low risk process at the University of Canterbury. Please do not hesitate to contact Dr McAuliffe if you have any concerns regarding your participation in this project (see contact details below). If you would like to speak with someone not involved in the study, please contact the Secretary of the Human Ethics Committee, University of Canterbury, Registry, Level 6.

FOR MORE INFORMATION

Should you have further questions regarding the research, please feel free to contact Joseph Dalrymple-Alford on 022 125 7118 or email jja34@uclive.ac.nz. Alternatively, Dr Megan McAuliffe on 364 2987 ext. 7075 or email megan.mcauliffe@canterbury.ac.nz.

Appendix 3: Speaker information sheet

INFORMATION SHEET – SPEAKER

1.04.2013

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Project Name:

You are invited to take part in the research project titled: "Does vocabulary knowledge influence speech recognition?

Please take the time to read this information sheet thoroughly and consider whether you would like to participate. Your participation is entirely voluntary (your choice). The following research team is conducting this study:

Principal Investigator:

Joseph Dalrymple-Alford, Master of Audiology Student

Associate Investigators:

Dr Megan McAuliffe, Associate Professor, Department of Communication Disorders

Dr Don Sinex, Senior Lecturer, Department of Communication Disorders

This study looks at how listeners comprehend speech in noise and how the task relates to vocabulary knowledge and memory. You will be asked to provide samples of your speech that will be used as part of a speech recognition experiment with older adult listeners. An understanding of how people understand speech will be helpful for audiologists in the development of treatment plans.

(1) You will be asked to read lists of short phrases. Approximately 40 people will hear your recorded phrases. Your participation will be required for one session of approximately 60 minutes. The recordings will be made on a stand alone recording system using a microphone in a noise-reduced room at the location noted below. There will be no identifying information included in the recording.

CONFIDENTIALITY

- Your privacy and confidentiality will be maintained at all times.
- All information will be kept in a locked filing cabinet at the Department of Communication Disorders, University of Canterbury. Only the researchers or research assistants involved in the project will have access to this information.
- The results of this project will be published; however, no material which could personally identify you will be used in any reports on this study.
- If you wish, you will be advised of the results of the study.

COMPENSATION FOR PARTICIPATION

You will receive a $20 petrol voucher for your participation in this study.

RISKS OF PARTICIPATION

There are no physical risks to participating in this project. However, if you feel uncomfortable or unable to continue at any time you can withdraw from the study.

LOCATION

The recording of speech samples will take place at the research laboratory at Room 801, Level 8, Rutherford Building, University of Canterbury (the Department of Communication Disorders research and postgraduate facility).

WITHDRAWING FROM THE STUDY

It is important to note that this study is voluntary and that you can withdraw from it at any time. This will in no way jeopardise any of your future dealings with the Department of Communication Disorders. If you choose to withdraw from the study, any data collected prior to withdrawal will not be used for research purposes without your consent.

ETHICS

This proposal has been reviewed and approved by the Department of Communication Disorders, University of Canterbury, and the University of Canterbury Human Ethics Committee Low Risk process.

Please do not hesitate to contact Dr McAuliffe if you have any concerns regarding your participation in this project (see contact details below). If you would like to speak with someone not involved in the study, please contact the Secretary of the Human Ethics Committee, University of Canterbury, Registry, Level 6.

FOR MORE INFORMATION

Should you have further questions regarding the research, please feel free to contact Joseph Dalrymple-Alford on 022 125 7118 or email jja34@uclive.ac.nz. Alternatively, Dr Megan McAuliffe on 364 2987 ext. 7075 or email megan.mcauliffe@canterbury.ac.nz.